

This gene is expressed primarily in keratinocytes and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, integumentary, or neurological and behavioural disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. brain, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or detection of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Alternatively, expression within keratinocytes indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia).

congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 10 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 535 of SEQ ID NO:73, b is an integer of 15 to 549, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

20 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

LCSTPVPTLFCPRIVLEVVLVRLSISEQCRRVSSQVTVASELRHRQWVERTLRSR
QRQNYLR (SEQ ID NO:366). Polynucleotides encoding these polypeptides are also encompassed by the invention.

25 This gene is expressed primarily in osteoclastoma.
Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal disorders, and diseases of the haemopoietic and immune system, particularly cancer. Similarly, polypeptides and antibodies directed to these
30 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bones, immune and haemopoietic system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.skeletal,hematopoietic, and cancerous and wounded tissues) or bodily fluids
35 (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard

gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:223 as residues: Ser-59 to Glu-67.

5 The tissue distribution in osteoclastoma tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the bones, immune and haemopoietic systems and cancer. Moreover, the protein may play a role as a therapeutic in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis, bone cancer, as well
10 as, disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders. For example, in rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia
15 congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
20 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
Accordingly, preferably excluded from the present invention are one or more
25 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 576 of SEQ ID NO:74, b is an integer of 15 to 590, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

When tested against dermal fibroblast cell lines, supernatants removed from
35 cells containing this gene activated the EGR1 (early growth response gene 1) promoter element. Thus, it is likely that this gene activates fibroblast cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell

types upon activation, leading the cells to undergo differentiation and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

ARGETAYDGAAVEFQEPLSSCLFSSLNPHHWPTLGVGRPVMLTLEDKD (SEQ
5 ID NO:367), ELLQCQMLEASTLIHLHHPRPGFPALCSFLGFRHHLHHDALCIRV
LPEDLEAKLCVSLHQQLHRLGLCLPGFGAACPGDQGSEDEARPPAVLRAVALLR
AGLRHLSVHSGWYHLPH SRNGLPLLALVVHFPEYGGGPREPVPVGQSG
EFGRRTELSTKGDTGDSRNShLAQDMAASLPFFKPCECTHV AVCSPPHPLCQ
YLCL (SEQ ID NO:368), LQCQMLEASTLIHLHHPRPGFPALCSFL (SEQ ID
10 NO:369), HQLLHRGLCLPGFGAACPGDQGSEDEARPPA (SEQ ID NO:370),
and/or LALVVHFPEYGGGPREPVPVGQSGEFGR (SEQ ID NO:371).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, male reproductive and endocrine disorders, cancer, particularly testicular cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive and endocrine systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, testes, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:224 as residues: Lys-53 to Leu-60, Pro-94 to Gln-99, Ser-176 to Gly-184, Ser-30 199 to Val-207.

The tissue distribution in testes, combined with the detected EGR1 biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of male reproductive and endocrine disorders, including aberrant testicular function (e.g. endocrine function, sperm maturation). Moreover, in light of the EGR1 activity, it may also be useful in the diagnosis and treatment of a variety of proliferative disorders, especially testicular cancer. Protein, as

well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 5 ID NO:75 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 10 a-b, where a is any integer between 1 to 1042 of SEQ ID NO:75, b is an integer of 15 to 1056, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 66

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: QSWTAPAAPARLPMALPQMQCDGSHLASTLRYC (SEQ ID NO:372), QSAAQWFWWPGRSASLGGAKGMQPPSLASWPXPRSIRCL RAPAPC 20 SXPSASSAAVQVACCCSLACCGPSRPASQGHRLWDPYHLSRDLYYLTVESSEK ESCRTPKVVDI PTYEEAVSFPVAEGPPTPPAYPTEEALEPSGRDALLSTQPA WPPPSYESISLALDAVSAETTPSATRSC SGLVQTARGGS (SEQ ID NO:373), GSTGLWRGDGRGPIEGGPGMLAL TDHSRVSFPVAEGPPTPPAYPTEEAL EPSGRDALLSSVXGASWPGWAVASPSLHQAKQSVPATRTTVPLTVM Q (SEQ 25 ID NO:374), QWFWWPGRSASLGGAKGMQPPSLASWP (SEQ ID NO:375), SSAA VQVACCCSLACCGPSRPASQGHRLW (SEQ ID NO:376), VSFPVAEGPPTPPAYP TEEALEPSGRDALLS (SEQ ID NO:377), and/or RVSFVVAEGPPTPPAYPTEE ALEPSG (SEQ ID NO:378). Polynucleotides encoding these polypeptides are also encompassed by the invention.

30 This gene is expressed primarily in pituitary gland.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine disorders, such as dwarfism. Similarly, polypeptides and 35 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this

gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. endocrine, immune, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in pituitary indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of disorders of the pituitary gland and endocrine system. Moreover, considering the vital importance of the pituitary in serving as a master regulator for various endocrine glands, the protein product of this gene would also be useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 916 of SEQ ID NO:76, b is an integer of 15 to 930, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

35 SNEILLSFPQNYYIQWLNGSLIHGLWNLASLFSNLCLFVLMPFAFFLESEGFA
GLKKGIRARILETLVM LLLLALLILGIVWVASALIDNDAAS (SEQ ID NO:379).

Polynucleotides encoding these polypeptides are also encompassed by the

invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

This gene is expressed primarily in the developing brain, liver and heart, and to a lesser extent, in cancerous tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, neural, hepatic, or cardiopulmonary and haemopoietic disorders, in addition to cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal tissues and the haemopoietic and neural systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developmental, neural, hematopoietic, hepatic, cardiovascular, pulmonary, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, bile, serum, pulmonary surfactant or sputum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:226 as residues: Glu-67 to Asn-74, Glu-88 to Asn-93, Lys-95 to Ser-105, Arg-152 to Ala-164, Ala-204 to Arg-210, Phe-254 to Thr-262, Pro-295 to His-311.

The tissue distribution in developing brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of haemopoietic and developmental diseases and cancers. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or

neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the relatively specific expression of this gene product during 5 embryogenesis indicates that it may be a key player in the proliferation, maintenance, and/or differentiation of various cell types during development. It may also act as a morphogen to control cell and tissue type specification. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers, which include, but are not limited to 10 the following tissues or cells: pulmonary, immune, neural, hematopoietic, or hepatic tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 15 ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 20 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 4449 of SEQ ID NO:77, b is an integer of 15 to 4463, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 68

The translation product of this gene shares sequence homology with a putative yeast transmembrane protein which may play an important role in intercellular signalling, intracellular transport, or regulation of cellular homeostasis. In specific 30 embodiments, polypeptides of the invention comprise the following amino acid sequence: PTRPVLLAINGVTECFTFAAMSKEEVDRYNFV (SEQ ID NO:380), and/or NDKLLFLKGFWSSLKNETPPPHFRLRMVTGVSCSGTLWCLISGV AVTPLQSPQWG SYTECVPPTELPIAGPGASGVQASLKSRSRHFVSASGHT (SEQ ID NO:381). Polynucleotides encoding these polypeptides are also encompassed by the 35 invention.

This gene is expressed primarily in pulmonary, immune cells, epididymus, and testis tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the reproductive organs, immune, and pulmonary systems, 5 in addition to endothelial and epithelial tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, respiratory and reproductive systems, expression of this gene at significantly higher or lower levels may be detected 10 in certain tissues or cell types (e.g.pulmonary, immune, reproductive, testes, epididymus, endothelial, epithelial, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, seminal fluid, pulmonary surfactant or sputum, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., 15 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:227 as residues: Arg-45 to Thr-52, Tyr-60 to Gly-66, Ala-87 to Trp-92, Leu-105 to Ser-115.

20 The tissue distribution and homology to putative transmembrane protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases of the reproductive, pulmonary and immune system. Moreover, the protein product of this gene may be useful in the diagnosis, treatment, and/or prevention of a variety of male reproductive disorders, which include. 25 but are not limited to, aberrant testicular function, male sterility, impotence, or related endocrine disorders. Protein may also serve a role as a contraceptive. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available 30 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 35 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 777 of SEQ ID NO:78. b is an integer of 15 to

791, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 69

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

10 SENRIYRNGLEKMRREVTIGRSSICLDQQVKAGNAVHHQWLKYVCWMVVVV
GGSGVGDG NLGM (SEQ ID NO:382). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in PMA induced T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, such as inflammatory or immunodeficiency conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, 20 particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene 25 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:228 as residues: Ser-62 to Thr-73, Phe-80 to Gln-88.

The tissue distribution in T-cells indicates that polynucleotides and polypeptides 30 corresponding to this gene are useful for study and diagnosis of immune system disorders. More specifically, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved 35 in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease,

sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1278 of SEQ ID NO:79, b is an integer of 15 to 1292, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

This gene is expressed primarily in monocytes. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, which include, but are not limited to, leukemias, lymphomas, AIDS, arthritis and asthma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard

gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in monocytes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, immunodeficiencies (e.g. AIDS), immuno-suppressive conditions (transplantation) and hematopoietic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Moreover, this gene may also be useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1269 of SEQ ID NO:80, b is an integer of 15 to 1283, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

When tested against dermal fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) promoter element. Thus, it is likely that this gene activates fibroblast cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-

STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

- 5 NWSGRRLRMWPSAALSPA VSSPALALTSPPKPLKGEVWLRWKLLGSRAVGLF
AF IALGTQSPLLHRACLPVRQSWGCSEHKAY PILRLQP DLETQVGPGHGVN
WDLRTQIRTIGELGGDGGCSE MRPLF (SEQ ID NO:383), and/or NLFSTPCKRQ
KLKLEWTEAPNVALRCSLSCSLIPGLSPDLSSEAPEGRSVAKMEIARQQSCWL
VCI YCFRNPESTLAPGLPACEAELGLLRAQGLPHPASPARLGN TGGAWPR
10 SKLGSQNTN (SEQ ID NO:384), SSPALALTSPPKPLKGEVWLRWKLLG (SEQ
ID NO:385), EHKAY PILRLQP DLETQVGPGHGVNWDL (SEQ ID NO:386), and/or
ALRCSLSCSLIPGLSPDLSSEAPEGRSV (SEQ ID NO:387). Polynucleotides
encoding these polypeptides are also encompassed by the invention. The gene encoding
the disclosed cDNA is believed to reside on chromosome 11. Accordingly,
15 polynucleotides related to this invention are useful as a marker in linkage analysis for
chromosome 11.

This gene is expressed primarily in placenta.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental anomalies, fetal deficiencies, pre-natal disorders and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the
25 reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, placental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,
30 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:230 as residues: Gly-22 to Gly-29, Gln-37 to Ala-44.

- The tissue distribution in placental tissue, combined with the detected EGR1
35 biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of developmental anomalies, fetal deficiencies and pre-natal disorders. In addition it may be useful in the detection and

treatment of ovarian and endometrial cancers. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 694 of SEQ ID NO:81, b is an integer of 15 to 708, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

20 LAPECCCCSVTYPRALVPRPCCPEPRAPLQLTLGLFSANPVNASPWGRCRSRR
GRGNLPLGHPVSTAFSSGDS (SEQ ID NO:388), and/or NTLHSKLVPSVYHSTE
KSCLV CFGMCPSIYKKMKSVLLIGTRMLLWLSHISQGPRPEAVLPRAPSP
SAAHPWLVFRKPGKRKPLGQMVKQK REGKPASGSPC (SEQ ID NO:389), YPR
ALVPRPCCPEPRAPLQLTLGLF (SEQ ID NO:390), and/or VLLIGTRMLL
25 WLSHISQGPRPEAVLPR (SEQ ID NO:391). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

30 This gene is expressed primarily in infant brain, and to a lesser extent, in placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and neurological disorders. Similarly, polypeptides and 35 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developmental and neurological

systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, developmental, neural, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an 5 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:231 as residues: Thr-45 to Arg-50.

10 The tissue distribution in fetal brain and placenta indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of various developmental and neurological disorders and diseases. The protein product of this gene is useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers 15 Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered 20 behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in 25 the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Expression within fetal tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other 30 proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

35 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:82 and may have been publicly available prior to conception of the present

invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of
5 a-b, where a is any integer between 1 to 1450 of SEQ ID NO:82, b is an integer of 15 to 1464, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:82, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 73

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

15 WIVVMFGKVLKIKDFMSTYSHTYTHTHMHAHTHTHTLTLQNVTLVAISDS
DK ALLIF (SEQ ID NO:392), MTLLIAEKTWRRPWPCQWGYLGAEGDRHLEG
RSLSLRHLQGAETPVLPNPDQLQLPSHIGKQAWSH AL GSL (SEQ ID NO:393),
MSTYSHTYTHTHMHAHTHTHTLTLQNVTLVAISDS
20 MSLSLRHLQGAET (SEQ ID NO:394), and/or GAEGDRHLEG
GRSLSLRHLQGAET (SEQ ID NO:395). Polynucleotides encoding these polypeptides are also encompassed by the invention.

25 This gene is expressed primarily in the spleen of patients with lymphocytic leukemia.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, lymphocytic leukemia and other cancers, as well as immune disorders such as AIDS, arthritis and asthma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at 30 significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spleen tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of

lymphocytic leukemia and other cancers, as well as other immune disorders and conditions including, AIDS, arthritis, asthma and microbial infection. Furthermore, the protein product of this gene may be useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia,

5 thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, 10 immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:83 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

20 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 602 of SEQ ID NO:83, b is an integer of 15 to 616, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:83, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

When tested against Jurket and fibroblast cell lines, supernatants removed from 30 cells containing this gene activated both the GAS (gamma activating sequence), and the EGR1 (early growth response gene 1) promoter elements. Thus, it is likely that this gene activates immune or fibroblast cells through the JAK-STAT and/or EGR1 signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal 35 transduction pathway involved in the differentiation and proliferation of cells.

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation

of cells. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid

5 sequence: VVEPGLKASLGA

MSTLFPSLFPRVTETLWFNLDRPCVEETELQQQEQQHQAWLQSIAEKDNNLVPI
GKPASEHYDDEEEEDD EDDEDSEEDSEDDEDMQDMDEMNDYNESPDDGEVN
EVDMEGNEQDQDQWMI (SEQ ID NO:396), LFPRVTETLWFNLDRPCVEETEL
(SEQ ID NO:397), and/or YNESPDDGEVNEVDMEGNEQDQQ (SEQ ID NO:398).

10 Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

15 This gene is expressed primarily in cells of the immune and haemopoietic systems, and to a lesser extent, in several other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and haemopoietic disorders, such as multiple myeloma, 20 immunodeficiencies, and inflammatory conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and haemopoietic systems, expression of this gene at significantly higher or lower levels may be detected in certain 25 tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:233 as residues: Pro-21 to Gly-30.

The tissue distribution in immune tissues and cells, combined with the detected GAS and EGR1 biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the 35 immune, haemopoietic, and integumentary systems. In addition, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia.

thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:84 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 914 of SEQ ID NO:84, b is an integer of 15 to 928, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:84, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: MGFDIHGVLGAEAVAEPREKKQE RAKWAPHDYDDPSLS LQDLLISWMISTWLIPMWKCQATIWFSLIQRLLNAYCMPGNFRHWEIAANTTN KT PGLMDFKFL (SEQ ID NO:399), EPREKKQERAKWAPHDYDDPSLSQLDL (SEQ ID NO:400), and/or MPGNFRHWEIAANTTNKT PGLMDF (SEQ ID NO:401). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on the X chromosome. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for the X chromosome.

This gene is expressed primarily in fetal liver and spleen, and to a lesser extent, in prostate cancer and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, reproductive, immune, and haemopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and developing systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developmental, hepatic, reproductive, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, bile, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in developing and immune tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of disorders of the haemopoietic and developing immune systems. In addition, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Moreover, the expression within fetal tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. The protein may also show utility in the treatment or diagnosis of various hepatic or reproductive disorders, which include, but are not limited to hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells, and prostate cancer, and/or congenital defects such as X-linked conditions. Protein, as well as, antibodies directed against the

protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 5 ID NO:85 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 10 a-b, where a is any integer between 1 to 709 of SEQ ID NO:85, b is an integer of 15 to 723, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:85, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 76

This gene is expressed primarily in fetal spleen and Wilm's tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, haemopoietic, immune, developmental, or renal disorders, such as congenital defects, multiple myeloma, or Wilm's tumor. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders 25 of the above tissues or cells, particularly of the haemopoietic and developing systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developmental, immune, hematopoietic, renal, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a 30 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the haemopoietic and developing systems and cancer. In addition, 35 polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the

production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, 5 immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility 10 in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or 15 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:86 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the 20 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 556 of SEQ ID NO:86, b is an integer of 15 to 25 570, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:86, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the (gamma activating sequence) promoter element. Thus, it is likely that this gene activates promyelocytic cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal 30 transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS 35

element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in induced T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells, combined with the detected GAS biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune and inflammatory diseases. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:87 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 625 of SEQ ID NO:87, b is an integer of 15 to 639, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:87, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 78

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: QSVPSPLAPPLPPSLPSFLFTETRSHYVARLVSNSWAQM ILLPWPLKVLGLDVSHCAWPKSVFLQAMEEIADFCLFSVKYQVSSMTCF DRT SYMKNTYL (SEQ ID NO:402), and/or LFTETRSHYVARLVSNSWAQMILLPWP (SEQ ID NO:403). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in bone marrow. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, anemias (leukemias), immune deficiencies and other hematopoietic-related disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in bone marrow indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of

hematopoietic and immune disorders, which include, but are not limited to the following: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immuno-suppressive conditions (transplantation) and other hematopoietic disorders, such as multiple myeloma. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:88 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 694 of SEQ ID NO:88, b is an integer of 15 to 15 708, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:88, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 79

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

SQIKSEKKHIGKAYTCTQTQSTGMQSTLTIVAKKSRNHTESYTRKKQENQIV
25 LIPWHQKKHPEGTHTCSHSLRRDTNTAADTQRKIRAHRYTYRRDKYSDTLVTH
DHYGDKHPSNTHTQPR XEFLQPGGSTNSRAAPRXSSFCPFS EGYS
SWGYH (SEQ ID NO:404), GMQSTLTIVAKKSRNHTESYTRKKQ (SEQ ID
NO:405), KKHPGTHTCSHSLRRDTNTAADT (SEQ ID NO:406), and/or RRDKY
30 SDTLVTHDHYGDKHPSNT (SEQ ID NO:407). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, such as leukemias, lymphomas, AIDS, arthritis and asthma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification

of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, 5 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID 10 NO:238 as residues: Asp-38 to Leu-43.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including leukemias, lymphomas, AIDS, arthritis and asthma, as well as other conditions which potentially implicate the immune system, such as 15 atherosclerosis, cancer and infection. In addition, This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an 20 agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune 25 infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies 30 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:89 and may have been publicly available prior to conception of the present 35 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 935 of SEQ ID NO:89, b is an integer of 15 to 949, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:89, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 80

In specific embodiments, polypeptides of the invention comprise the following 10 amino acid sequence: KHLPLKAPIDLDNKNSCMFCSDIFCRFH HSTAWLFL GRITDRILGLHHYLIRYQFEIENLCLMKIVIPVVSMTNCQFDLQLKQNLHYH (SEQ ID NO:408), APIDLNDKNSCMFCSDIFCR (SEQ ID NO:410), and/or IENLCLMKIVIPVVSMTNCQFDLQL (SEQ ID NO:409). Polynucleotides encoding these polypeptides are also encompassed by the invention.

15 This gene is expressed primarily in prostate carcinoma cell line stimulated with 30 nM synthetic androgen, R1881 cells and, to a lesser extent, in activated monocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as 20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive or immune disorders, particularly prostate cancer and 25 prostate ailments. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate, expression of this gene at significantly higher or lower 30 levels may be detected in certain tissues or cell types (e.g. immune, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in the prostate indicates that polynucleotides and 35 polypeptides corresponding to this gene are useful for the diagnosis and intervention of prostate cancer and prostate ailments, or related proliferative conditions in either said tissue or other tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:90 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1157 of SEQ ID NO:90, b is an integer of 15 to 1171, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:90, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

The translation product of this gene shares strong sequence homology with human protocadherin 42 (GenBank accession no. gil387675), PCDH7 (BH-Pcdh)a, and its associated isoforms PCDH7 (BH-Pcdh)b, and PCDH7 (BH-Pcdh)c which are thought to be important in tissue and cell-cell adhesion, repair and development (See Genbank Accession Nos.gnl|PID|d1026122 (AB006755), gnl|PID|d1026123 (AB006756), and gnl|PID|d1026124 (AB006757)). The polynucleotides encoding this gene have been gened by another group subsequent to our filing (See Yoshida K, et al. Genomics 1998 May 1;49(3):458-61, which is hereby incorporated by reference). The cytoplasmic domain of cadherin interacts with the cytoskeleton through catenins and other cytoskeleton associated proteins. The cytoplasmic domain is not present in all cadherins, but in those which possess it, it is essential for the cadherins adhesive function. The cadherins which do not possess a cytoplasmic domain appear to function via a different method from those with a cytoplasmic domain. This protein sequence is involved in cell-cell adhesion. This sequence may have regulatory functions in the cell, as well as the cell-cell adhesive properties. Antibodies produced against this sequence are useful for modulating the binding activity of protocadherins, and can be used therapeutically. BH-Pcdh has an extracellular domain consisting of seven repeats of the cadherin motif (EC 1 to 7). EC2 of BH-Pcdh is unique in having a 55-amino-acid insertion in the middle of the motif. There are three isoforms of BH-Pcdh, denoted -a, -b, and -c, which have different cytoplasmic tails and a 47-amino-acid deletion in the EC2-3 region of BH-Pcdh-c. While only a 9.0-kb message was detected in normal tissues, 4.5- and 9.0-kb mRNA species were seen in the human lung carcinoma cell line A549. Furthermore, only the 4.5-kb mRNA was detected in HeLa cell S3 and

human gastric cancer cell lines MKN28 and KATO-III. Southern blot analysis indicated that the BH-Pcdh gene is likely to be conserved among various vertebrates. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

- 5 GTSVNESVSNATAIDSQIARSLHIPLTQDIAGDPSYEISKQRLSIVIGVVAGI (SEQ ID NO:411), PKIKMAMKPAKKITKTFLHPNSMTNLKSLKRTRKTKNLSSLSTA
ALSLWRLLSQMDRGMIVSMRSCQTAQ AWGDTGPLMVGPAPLTWQGITNL
VPHCLLFSFIPSHQLQEKNTRPYKIYHQPTHLWEQETTFQLDQITAL STAVKP
ITSTANRCVYIHTLLCLAEFHSNMMHLHYAPYCDDLSTPKPAGACPWPWGVSQS
10 LLVPLVVHFIF ESFSFSYTEK (SEQ ID NO:412), CSIMHHTVMTFLRNLLLEPA
LGRGVSAZHCLFHLLYILFL SLFLSHIQKNSMKIK (SEQ ID NO:413), TAIDS
QIARSLHIPLTQDIAGDPSYEISK (SEQ ID NO:414), YCRSKNKNGYEAGKKDH
EDFF (SEQ ID NO:415), GPGSPDLARHYKSSPLPTVQ (SEQ ID NO:416), and/or
15 LPPANTFVGAGDNISIGSDHCSEYS (SEQ ID NO:417). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed primarily in ovarian tumors, and to a lesser extent in, striatum and HL-60 cells.

- 20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and reproductive dysfunction, in addition to cardiovascular and neural disorders, such as atherosclerosis, and neurodegenerative disorders, such as
25 Alzheimer's and Parkinson's, or other disorders resulting from aberrant cell-adhesion. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, nervous and immune systems, expression of this gene at significantly
30 higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, neural, cardiovascular, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
35 individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:240 as residues: Tyr-15 to Leu-59, Ala-68 to Asp-85, Pro-87 to Asn-96, His-120 to Lys-129, Ser-153 to Gln-170.

The tissue distribution in ovarian and muscle tissue, combined with the strong homology to various cadherins indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, study and treatment of various neoplastic disorders such as squamous cell carcinomas and related tumors, and nervous system and reproductive disorders. Considering the vital importance of cell-adhesion amongst various cellular functions, in particular chemotaxis by the immune and hematopoietic cells indicates that this gene product may play a direct, or in-direct role in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also play an in-direct role in the regulation of a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-

inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against 5 the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:91 and may have been publicly available prior to conception of the present 10 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 15 a-b, where a is any integer between 1 to 1137 of SEQ ID NO:91, b is an integer of 15 to 1151, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:91, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

20

The translation product of this gene shares sequence homology with the G-protein coupled receptor TM3 consensus polypeptide which may implicate an important function for this protein in various signal transduction pathways. G-protein coupled receptors are known to have a variety of functions including modulating immune 25 system tissue through interaction with cytokines and lymphokines. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

GTSNASVSPTICICMCGYVHIWFFICLCVYLKVLQGSACPWIAAAVVMRRMRKVQEKG
EVFRNMAATWAL RSGIQSLNSLVSSAFFTIFMTLGSSWNLIVSLSSLV
30 NWTGLFSFYFSRN (SEQ ID NO:418), CLCVYLKVLQGSACPWIAAAVV
MRRMRK (SEQ ID NO:419), and/or TIFMTLGSSWNLIVSLSSLVNWTGLF (SEQ
ID NO:420). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in breast lymph node. 35 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, breast cancer, or other immune or reproductive disorders and diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the
5 immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, reproductive, breast, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, breast milk, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the
10 expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:241 as residues: Cys-34 to Gly-48.

The tissue distribution in breast lymph nodes and homology to a conserved G-
15 protein coupled receptor TM3 consensus sequence indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment for breast cancer or immune diseases. Considering the vast roles which G-protein coupled receptors play in the maintenance of important cellular fuctions, the secreted protein may have a very wide range of biological acitivities. Typical of these are cytokine, cell
20 proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating
25 wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation
30 of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. Protein, as well as, antibodies directed against the protein may show
35 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:92 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more

- 5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 700 of SEQ ID NO:92, b is an integer of 15 to 714, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:92, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 83

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

- 15 QPDIPVLPVGFSQNCASFKVSGCWKGGLIAEKVGTLGTPKGRR AWPETEF
FRFLEPGLP (SEQ ID NO:421), and/or RGFRMAQPLVNTFQVAVPVEDL
APQQNPSRFPADPALLSFLTG SILAPGKVIWVNVSFTAIIWPTWDSMAI
GELTIASHASMTLHIGRPGSRKRKNVSUGHARLPFGVPSVPT FSAISPP
FQQPETLKEQF (SEQ ID NO:422), EDLAPQQNPSRFPADPALLSFLTG (SEQ ID
20 NO:423), and/or TWDSMAIGELTIASHASMTLHIGRPGSRK (SEQ ID NO:424).
Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in activated T-cells, hepatocellular tumor, pancreas islet cell tumors, and hemangiopericytoma.

- 25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hepatic, and endocrine disorders, such as cancers, particularly of T-cells, hepatocellular tumors and pancreas islet cell tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hepatic, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or
30 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:242 as residues: Glu-43 to Lys-50, Ser-53 to Phe-60.

- The tissue distribution in T-cells, hepatocellular tumors, and pancreatic islet cell tumors indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of immune, hepatic, and endocrine disorders, and other cancer types. Expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders in various tissues, aside from those disclosed above. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.
- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:93 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 796 of SEQ ID NO:93, b is an integer of 15 to 810, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:93, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 84

- In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:
- VSPQLMGIKREPSAAQLSVGEEHTLDREGRELVDLPGQPSQKIKIKNKSSLHPG
LIIPP AHYKTATTNLF (SEQ ID NO:425), and/or PSAAQLSVGEEHTLDREGREL (SEQ ID NO:426). Polynucleotides encoding these polypeptides are also encompassed by the invention.

35

This gene is expressed primarily in hepatocellular tumors. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hepatic disorders, such as liver diseases and hepatocellular tumor, including proliferative disorders in other tissues and cell types. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. hepatic, proliferating, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in hepatocellular tumor tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of hepatocellular tumor or other liver disorders. Specifically, polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:94 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1162 of SEQ ID NO:94, b is an integer of 15 to 1176, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:94, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 85

35

When tested against reh cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is

likely that this gene activates B-cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

- 5 Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: NCDHDFIQPLHTPMMSAL FQSEFS (SEQ ID NO:427), SILNM GLFTEQRWPAAARCARQSTVAGAIRRARGTVTMWQVAGAAW ASPDRRAKV
10 HPCRHAAPCLPSPCRRGLQMSGPLQATRGRVTLSHQVGCKRATGSIENSL (SEQ ID NO:428), QKSKGSPLQTCCSLPTLPMQERPADEWSTPGDQGKSYIK KPPGGLQKGHRLHRKLTLKQGRHRGVE GLNEIMVTVLKEEFPVSKPGLNV LPTFHRHHECYQHGMNLTARISVVS (SEQ ID NO:429). ARQSTVAGAIRR ARGTVTMWQVAGA (SEQ ID NO:430), PCRRGLQMSGPLQATRGRVTLSHQ (SEQ ID NO:431), LPMQERPADEWSTPGDQGKSYIKKPP (SEQ ID NO:432), and/or NVLPTFHRHHECYQHGMNLTARI (SEQ ID NO:433). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human fetal kidney, adult testis, T-cell lymphoma, and a fetal liver/spleen cDNA library.

- 20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal, developmental, reproductive, immune, or hematopoietic disorders, particularly kidney disease, lymphoma, congenital defects, multiple myeloma, SCID, male sterility, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney and immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, reproductive, renal, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 30 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:244 as residues: Gly-35 to Gly-40.

The tissue distribution in fetal kidney and T-cells, combined with the detected GAS biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis or treatment of kidney diseases or immune disorders, especially cancers. Specifically, this gene or gene product could be used in
5 the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Expression within
10 fetal tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly,
15 developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
20 ID NO:95 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
Accordingly, preferably excluded from the present invention are one or more
25 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1014 of SEQ ID NO:95, b is an integer of 15 to 1028, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:95, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 86

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates promyelocytic cells through the JAK-STAT
35 signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in breast, human embryo, and chronic spleen
5 lymphocytic leukemia.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, developmental, hematopoietic or immune disorders, such 10 as breast cancer, congenital birth defects, or leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast or hematopoietic systems, expression of this gene at significantly higher or lower levels may be detected in certain 15 tissues or cell types (e.g. reproductive, immune, hematopoietic, developmental, breast, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an 20 individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:245 as residues: His-2 to Asn-8, Gln-35 to Phe-44.

The tissue distribution in breast and lymphocytic leukemia cells, combined with the detected GAS biological activity indicates that polynucleotides and polypeptides 25 corresponding to this gene are useful for the diagnosis or intervention of breast cancer, leukemia or other hematopoietic related disorders. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of 30 cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the 35 expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies

directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:96 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 733 of SEQ ID NO:96, b is an integer of 15 to 747, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:96, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 87

This gene is expressed primarily in brain containing medulla blastoma.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural disorders, particularly specific brain tumors such as medulla blastoma, as well as other diseases and conditions of the brain, such as schizophrenia, Alzheimer's disease, Tourette's syndrome, Parkinson's disease, Huntington's disease, mania, dementia, paranoia, depressive and addictive predispositions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution in brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of specific brain tumors such as medulla blastoma. In addition it may also be useful for the diagnosis and treatment of developmental, degenerative and behavioral conditions of the

brain and nervous system, such as schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, Tourette's syndrome, mania, dementia, paranoia, addictive behavior, obsessive-compulsive and sleep disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or 5 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:97 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the 10 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 614 of SEQ ID NO:97, b is an integer of 15 to 628, where both a and b correspond to the positions of nucleotide residues shown in 15 SEQ ID NO:97, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 88

20 When tested against Jurket cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates T-cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal 25 transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: INVLYCSRDSLGMGRTIMESSDYIKKGANVSPVLGVRQQ AV 30 (SEQ ID NO:434). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in adrenal gland tumor and T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 35 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the endocrine and immune or haemopoietic systems, particularly inflammatory or immunodeficiency conditions, such as AIDS. Similarly,

polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may 5 be detected in certain tissues or cell types (e.g. immune, hematopoietic, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the 10 disorder.

The tissue distribution in T-cells and adrenal gland tissues, combined with the detected GAS biological activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune and endocrine systems and cancer. Moreover, the secreted protein can also be 15 used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for 20 treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, 25 tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies 30 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:98 and may have been publicly available prior to conception of the present 35 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 890 of SEQ ID NO:98, b is an integer of 15 to 904, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:98, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 89

In specific embodiments, polypeptides of the invention comprise the following
10 amino acid sequence: SLLMYFVFKIFFQSLCVLGYCILPLTVA (SEQ ID NO:435).
Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, immune disorders. Similarly, polypeptides and antibodies directed to
these polypeptides are useful in providing immunological probes for differential
identification of the tissue(s) or cell type(s). For a number of disorders of the above
tissues or cells, particularly of the immune system, expression of this gene at
20 significantly higher or lower levels may be detected in certain tissues or cell types
(e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum,
plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
from an individual having such a disorder, relative to the standard gene expression
level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
25 having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
NO:248 as residues: Thr-43 to Thr-48.

The tissue distribution in dendritic cells indicates that polynucleotides and
30 polypeptides corresponding to this gene are useful for the treatment and diagnosis of
immune system disorders. In addition, polynucleotides and polypeptides corresponding
to this gene are useful for the treatment and diagnosis of hematopoietic related disorders
such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal
cells are important in the production of cells of hematopoietic lineages. The uses include
bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow
35 reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also
be involved in lymphopoiesis, therefore, it can be used in immune disorders such as
infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product

may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:99 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 10 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 562 of SEQ ID NO:99, b is an integer of 15 to 576, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:99, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 90

- In specific embodiments, polypeptides of the invention comprise the following 20 amino acid sequence:

RLWMTKAHPALRHLLLFTLALTLLAQGCCAVAPSGCADLAGFCSLGHS C
(SEQ ID NO:436). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human stomach.

- 25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, digestive and gastrointestinal conditions, particularly ulcers and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in 30 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.gastrointestinal, metabolic, mucosal, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, chyme, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken 35 from an individual having such a disorder, relative to the standard gene expression

level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:249 as residues: Pro-32 to Gly-38.

5 The tissue distribution in stomach tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study and treatment of gastrointestinal disorders, or other disorders afflicting mucosal or endothelial tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:100 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

15 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 699 of SEQ ID NO:100, b is an integer of 15 to 713, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:100, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 91

The translation product of this gene was found to have homology to the conserved K07F5.14 protein from *Caenorhabditis elegans* (See Genbank Accession No gnl|PIDle233697) which may be important in regulation of important cellular functions, including homeostasis and cell division. When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) pathway. Thus, it is likely that this gene activates promyelocytic cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

RTCTPWMGFWCLVCSLFAPVPTSRKYLVSKPGCYQRRRV FGVCFTKPL (SEQ

ID NO:437), WLLSEKKG (SEQ ID NO:438), and/or GVFYKAAVIG (SEQ ID NO:439). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in bone marrow and T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly multiple myeloma, immunodeficiencies, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in bone marrow and T-cells, combined with the detected GAS biological activity in U937 cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of immune and hormonal disorders and neoplasias. Specifically, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Moreover, the protein may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune

infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the
5 differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
10 ID NO:101 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of
15 a-b, where a is any integer between 1 to 635 of SEQ ID NO:101, b is an integer of 15 to 649, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:101, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 92

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

CKTSPLPKEGQSAVSVPVSSHFLAHSAPLSGGHAHVFARDGATGL (SEQ ID
25 NO:440), LGRGSGERKTPVSCFAQISKSRGGRSKSLTHLCTHTHTQVTEL
DVRMSHGCLRQHAGRLAPPPLRFCL TACWGRGEAETVWKDPASSQ
HPPPSEKPHRQDRHPERWHQPGGPIPGKHMRVSPGQRGRVCQEMGRNRM
(SEQ ID NO:441), FCLRDFKIWRGRLEAGRTEGRL AGERFGGEEDPSFLFC
SDFKVEGWAFEISHSLVHTHTHTGHGAGRADVTRVPAGTARWEAGSPTPSPV
30 LF DSLLGAAGRG (SEQ ID NO:442), AQISKSRGGRSKSLTHLCTHTHTQVTEL
(SEQ ID NO:443), EKPHRQDRHPERWHQPGGPIPGKHMR (SEQ ID NO:444),
GRLEAGRTEGRL AGERFGGEEDPSFL (SEQ ID NO:445), and/or
VTRVPAGTARWEAGSPTPSPVLF (SEQ ID NO:446). Polynucleotides encoding
these polypeptides are also encompassed by the invention. The gene encoding the
35 disclosed cDNA is believed to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in ovary, spinal cord, and fetal spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, reproductive, and neurological conditions. Similarly, 5 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and reproductive systems, expression of this gene at significantly higher or lower levels 10 may be detected in certain tissues or cell types (e.g. developmental, reproductive, ovarian, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily 15 fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:251 as residues: Pro-34 to Pro-53.

The tissue distribution in spinal cord and fetal spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study 20 and treatment of neural, hematopoietic, and developmental disorders. Specifically, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, 25 peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this 30 gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival.Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, 35 sexually-linked disorders, or disorders of the cardiovascular system. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia.

- leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include, but are not limited to bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. Protein, as well as,
- 5 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:102 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 683 of SEQ ID NO:102, b is an integer of 15 to 697, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:102, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 93

20

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: DEGVQGERLFRILRINGEKPYNFVVDYFHCEY (SEQ ID NO:447), KVVRIDNGILCSHKKTEIMSLQQHGIWRPYLKQTNTGTENQIPHTL TYKWELNFEYIXTQXRGXXDSEAYLKVEGGRREGIQLPIRYYVYLGDKIIC 25 SSSCSMHLLM (SEQ ID NO:448), HKDTCMSMFT AALFTIAKTWN (SEQ ID NO:449), MPINDRLDFKRWYV (SEQ ID NO:450), TMESYVAIKRQRSCPCSNM VGGGGHILSKLTQEQQTKYHILS LISGS (SEQ ID NO:451), EIMSLQQHGIW RPYLKQTNTGTEN (SEQ ID NO:452), and/or RREGIQLPIRYYVYLGDKIIC (SEQ ID NO:453). Polynucleotides encoding these polypeptides are also encompassed by the invention.

30

This gene is expressed primarily in bladder tissue from a human male.

35

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal, urogenital, and nephrotic conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a

number of disorders of the above tissues or cells, particularly of the gastrointestinal and excretory systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. renal, bladder, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:252 as residues: Arg-52 to Ala-57, Pro-66 to Thr-72.

The tissue distribution in bladder tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of gastrointestinal and urinary tract disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:103 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1274 of SEQ ID NO:103, b is an integer of 15 to 1288, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:103, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 94

This gene is expressed primarily in bladder tissue from a human male.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal, renal, and urinary tract conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the intestinal and urinary tract, expression of this gene at significantly higher or lower levels may be

detected in certain tissues or cell types (e.g. renal, urogenital, bladder, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in 5 healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in bladder tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of urinary tract and gastrointestinal disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above 10 listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:104 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the 15 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1013 of SEQ ID NO:104, b is an integer of 15 to 1027, where both a and b correspond to the positions of nucleotide residues shown 20 in SEQ ID NO:104, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 95

25 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: LHGEQVPI YIFLLMQPLNFECISFLNCIEQYSVGVI HNSV TIYACDREENCMDIRYL (SEQ ID NO:454), and/or GTSWASRFFTCH (SEQ ID NO:455). Polynucleotides encoding these polypeptides are also encompassed by the invention.

30 This gene is expressed primarily in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and inflammatory disorders, particularly immunodeficiencies 35 such as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of

the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:254 as residues: Lys-28 to Thr-34.

The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune system. Moreover, This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:105 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 696 of SEQ ID NO:105, b is an integer of 15 to 710, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:105, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 96

5 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

GPPRXFXPKKAILGXPPXGRVPPFRYRSRNSRGRPHXSAPRVRFCLENSWLR
(SEQ ID NO:456), and/or PLNTMMCMMCKMKVSPKIFSKLKRKYLNSNLTKL
EMQTVHLESSLASCSPNKGXVGRTR GVDPGNSGTGT (SEQ ID NO:457).

10 Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in lymphoma and frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and haemopoietic diseases, particularly neurodegenerative conditions such as Alzheimers and Parkinsons. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neural systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.. the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in frontal cortex and lymphoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases of the neural and haemopoietic systems. Specifically, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this

gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Moreover, the expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Since, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:106 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 516 of SEQ ID NO:106, b is an integer of 15 to 530, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:106, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 97

30

This gene is expressed primarily in the spleen of a patient with metastatic melanoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly metastatic melanoma and other cancers, as well as immune disorders and conditions such as anemias, AIDS.

arthritis and asthma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or 5 lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an 10 individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:256 as residues: Pro-26 to Asn-34.

The tissue distribution in spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of metastatic 15 melanomas and other cancers, as well as other immune disorders and conditions including leukemias, lymphomas, AIDS, arthritis, asthma and microbial infection. Furthermore, polynucleotides and polypeptides corresponding to this gene are useful 20 for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, or thrombocytopenia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex 25 vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies 30 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available 35 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:107 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 378 of SEQ ID NO:107, b is an integer of 15

to 392, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:107, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 98

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: GTVTQKRK CVFGKYLLSTCSLMFSSMHGACSWKA KQTSSSAGFLCLHVLCALQLTREKYKTWPWPSFI (SEQ ID NO:458), and/or 10 MKEGQGHVLYF SRVNCKAGHXTCRQRKPADELVCFAFQEQA PCILLNI RLQVLNKYLPNTHFLFCVTP (SEQ ID NO:459). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 8. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 8.

15 This gene is expressed primarily in pineal gland and synovial sarcoma. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine or skeletal disorders, including cancers. Similarly, 20 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.endocrine, pineal, skeletal, and cancerous and wounded 25 tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in pineal gland indicates that polynucleotides and 30 polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the endocrine system. In addition, polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), 35 adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-,hypoparathyroidism) , hypothalamus, and testes. Alternatively, the expression of this gene product in synovium would suggest a

role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders
5 such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
10 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:108 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
15 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 977 of SEQ ID NO:108, b is an integer of 15 to 991, where both a and b correspond to the positions of nucleotide residues shown in
20 SEQ ID NO:108, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 99

25 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:
TMTGIDSSPSEEILRQVGCKQQQKGVEHVEGSSAEAGEAARGGGAK GGGG
AAGKGTSKVGTLLRRTRGST (SEQ ID NO:460). Polynucleotides encoding these polypeptides are also encompassed by the invention.

30 This gene is expressed primarily in breast and fetal spleen. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the reproductive system and developing organs, particularly
35 congenital defects afflicting the immune or hematopoietic system, such as immunodeficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification

of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing and reproductive systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, developing, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:258 as residues: Gly-23 to Asn-30, Ser-37 to Asn-43.

The tissue distribution in fetal spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases involving developmental tissues and reproductive organs. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:109 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 898 of SEQ ID NO:109, b is an integer of 15 to 912, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:109, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 100

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

10 AQREAGSRP RRRKSLKAVAMLXVEMGGGCRGSMGP GPGYSAGSRVCRGSSL
PQVAPFNPSRAHLLPPPVG GGLNSVWLSGVQLSTPPYADWEGVGQSPQ
PRGPWMGSSSLGTVGPGCVLSCPTVKANGGSPCSEMLGER RLLEPSVG
PVSGCPERREGGHGARGAAGVVVKGHASVQLNFLSLI (SEQ ID NO:461),
KAEFTFAKEKNAKAQLGKKGTRWKHDKRKEIQLYGCVTLNDDPSCPPCPVP
15 TPPFWTA TYGSHGRFQKPPFSQHLRAGGAPVGLDCGAPTQYAARPHGPK
(SEQ ID NO:462), GCRGSMGPGPGYSAGSRVCRGSSLPQ (SEQ ID NO:463),
QPRGPWMGSSSLGTVGPGCVLS (SEQ ID NO:464), and/or GAAGVVVKGH
ASVQLNFLSLI (SEQ ID NO:465). Polynucleotides encoding these polypeptides are
also encompassed by the invention.

20 This gene is expressed primarily in endothelial, immune, and cancer cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases involving immune, endothelial, and haemopoietic tissues or 25 cells, particularly cancers, inflammatory or immunodeficiency conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, haemopoietic and endothelial systems, expression of this gene at significantly higher or 30 lower levels may be detected in certain tissues or cell types (e.g. endothelial, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an 35 individual not having the disorder.

The tissue distribution in immune and hematopoietic tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis

and treatment of disorders of the immune and haemopoietic systems, including cancer. More specifically, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is
5 expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell
10 mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in
15 the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
20 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:110 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more
25 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 861 of SEQ ID NO:110, b is an integer of 15 to 875, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:110, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 101

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: GKPLSAIFPICH MMFLPGKFNLGISHRCCRMT SPWDK
35 RQQLRQECKSDPHVQNPRIHFPESKNSFPSAYIFVSEGNGVSPSK WHCIY
SGTSLSH (SEQ ID NO:466), and/or GERGRYQSKYSATWMVTPHYLQTQRC

KLREMNSWIQGNEFLDSEHEGQIYIPVSIVDAYPKD (SEQ ID NO:467).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human kidney, and to a lesser extent, in liver.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, kidney, urogenital, hepatic, and endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
10 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal or endocrine systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. urogenital, kidney, endocrine, hepatic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, bile,
15 urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
20 NO:260 as residues: Glu-38 to Lys-43.

The tissue distribution in kidney indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of renal disorders, including noninflammatory and inflammatory lesions, and tumors of the kidney. Moreover, this gene or gene product could be used in the treatment and/or detection of
25 kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Alternatively, expression within liver
30 indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets
35 for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:111 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 445 of SEQ ID NO:111, b is an integer of 15 to 459, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:111, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 102

This gene is expressed primarily in kidney cortex and fetal tissue.utility_

The tissue distribution in kidney indicates that this gene or gene product could 15 be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's 20 syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 25 ID NO:112 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 30 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 595 of SEQ ID NO:112, b is an integer of 15 to 609, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:112, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 103

35

This gene is expressed primarily in ovary and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive and neurological conditions, particularly proliferative disorders, such as ovarian cysts or cancer, in addition to neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, reproductive, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in ovarian tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of reproductive disorders, such as infertility. Alternatively, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:113 and may have been publicly available prior to conception of the present

invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 5 a-b, where a is any integer between 1 to 1390 of SEQ ID NO:113, b is an integer of 15 to 1404, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:113, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 104

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ISIRGRIL

YK MAY FK VCVII WF QQ FC VEET SI IKN VRML TSEF QNS YAT PV SGLL PG AV AWR
15 GG AV YG WVR HAM QVL Q KEPT QPSS FLPP SD AASF FW GP ES RL HL TW (SEQ ID NO:468), KP FA FS AR NF PT ML SEAY FQ D P RMR QH HL GVER M TV AW VP SA IP AWR AS PRT QHH PSKP QHQ EGAQ KQG WHM NSG ILMSA YEHFL (SEQ ID NO:469), and/or HSK QNIC REV N IL KMFL HEIK KT VT DNIST QRR FT YNH QPGS VS IF SV TDILD FEVP FGL (SEQ ID NO:470). Polynucleotides encoding these 20 polypeptides are also encompassed by the invention.

This gene is expressed primarily in melanocytes, and PHA stimulated T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 25 not limited to, immune or integumentary system disorders, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain 30 tissues or cell types (e.g. integumentary, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of cancers and immune system disorders. Alternatively, the expression in melanocytes

indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:114 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 839 of SEQ ID NO:114, b is an integer of 15 to 853, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:114, and where b is greater than or equal to a + 14.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 105

This gene is expressed primarily in B cell lymphoma, and to a lesser extent, in dermal fibrosarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or integumentary disorders, particularly lymphatic and soft tissue cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in B-cell lymphoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the protein product of this gene may also be useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to

viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athletes foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma,
5 tendonitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed
10 against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:115 and may have been publicly available prior to conception of the present
15 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 831 of SEQ ID NO:115, b is an integer of 15
20 to 845, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:115, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 106

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element.
25 Thus, it is likely that this gene activates promyelocytic cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.
30 Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:
35

KVIDVIFSLPPGRKATFSCPLAPLSGAXGLPGGGANRPGPFLPCIQPWGPLRLPEGC (SEQ ID NO:471), MSSSLCPQGGKPPSLAPWPLCQGPXVCRVGVP

GLALSSPASSHGGLCDRKVAWLVPGPQARG RAAWFYFYLTLSVL (SEQ ID NO:472), and/or LALSSPASSHGGLCDRKVAWLVPGP (SEQ ID NO:473). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in T cells, fetal liver, and to a lesser extent, in various normal and transformed tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, or developmental disorders, including immunodeficiencies and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:265 as residues: Arg-5 to Pro-12.

The tissue distribution in B-cells and fetal liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of immune and developmental disorders. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. In addition, expression within fetal tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues

rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:116 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 746 of SEQ ID NO:116, b is an integer of 15 to 760, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:116, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 107

One embodiment of this gene comprises the following amino acid sequence:
MQRERWARPWMASTVESRMPEGKWRRFSTDLATWGATPARSWTKASRGSTT
AWTRLPMRSTMVLDKQERKQRSLAMGSTLLDRPGRKQTKRSKGSTLGSTRL
GRKQRNLAKGSTMLLTRLERXWRSLAQVPTMLLARPGRSCRMLIMGSTKPAR
RPTSC (SEQ ID NO:474). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in keratinocytes and tissues undergoing wound healing, and to a lesser extent, in osteoblasts and smooth muscle.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skin disorders; fibrosis; scarring; osteoporosis; osteopetrosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, bone, or connective tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. skin, bone, connective tissues, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:266 as residues: Gly-76 to Leu-83, Ala-108 to Glu-113, Ala-126 to Lys-132, Gly-145 to Leu-151.

10 The tissue distribution in keratinocytes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of skin disorders. Elevated expression of this protein in skin and keratinocytes suggest that it may be involved in keratinocyte proliferation, survival, and/or differentiation. Thus, it may play a role in such processes as fibrosis and wound healing. Similarly, expression of this protein in osteoblasts indicates that it may also play a role in osteoblast survival, proliferation, and/or differentiation, and that it may be 15 useful in the treatment of such disorders as osteoporosis or osteopetrosis.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:117 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the 20 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 974 of SEQ ID NO:117, b is an integer of 15 to 988, where both a and b correspond to the positions of nucleotide residues shown in 25 SEQ ID NO:117, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 108

30 The translation sequence of this gene shares homology with a mouse camodulin binding protein. The calcium-binding regulatory protein calmodulin is an essential subunit of the erythrocyte and other plasma membrane calcium ATPases. A rise in cytosolic calcium induces the binding of calcium ions to calmodulin, which triggers an allosteric activation of the calcium ATPase, and subsequently an export of calcium ions 35 from the cell is accelerated.

This gene is expressed primarily in teratocarcinoma cells, and to a lesser extent, in myeloid progenitor cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental defects, calcium-transport defects, in addition to immune or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of embryonic and fetal tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developing tissues, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:267 as residues: Tyr-124 to Gly-129.

The tissue distribution in teratocarcinoma cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of developmental defects as well as for organ regeneration. Moreover, expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Alternatively, the homology of the translation product of this gene to a mouse calmodulin binding protein indicates that the translation product of this gene may be useful for disorders involving calcium transport across the plasma membrane, for example. It has further been suggested this type of disorder may be responsible for disorders such as hypertension.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:118 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1933 of SEQ ID NO:118, b is an integer of 15 to 1947, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:118, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 109

One embodiment of this gene comprises polypeptides of the following amino acid sequence:

10 MRPLLGLLVFAGCTFALYLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAELREL
SEVI.REYRKEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWI.GLLI.
CCVLTSGATCCYLLSSIFGKQLVVSYFPDKVALLQRKVEENRNSLFFFLLFLR
LFPMTPNWFLNLSAPILNIPIVQFFFSVLIGLI PYNFICVQTGSILSTLTSLDA
15 LFSWDTVFKLLAIAMVALIPGTLIKFSQKHLQLNETSTANHIHSRKDT (SEQ ID
NO:475). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in ovarian tumor, and to a lesser extent, in smooth muscle and breast cancer.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, particularly of the ovary, musculature, and breast, such as rhabdomyosarcomas or fibroids. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. ovaries, breast, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, 25 breast milk, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID 30 NO:268 as residues: Arg-24 to Arg-29.

The tissue distribution in ovarian tumor tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of

cancer, particularly ovarian and breast cancers. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
5 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:119 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
Accordingly, preferably excluded from the present invention are one or more
10 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1434 of SEQ ID NO:119, b is an integer of 15 to 1448, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:119, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 110

The translation product of this gene shares sequence homology with bovine acrosin inhibitors IIa and IIb which is thought to be important as protease inhibitors.

20 This gene is expressed primarily in keratinocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, integumentary disorders, such as psoriasis, and wound healing abberations. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the integumental system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.integumentary, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:269 as residues: Tyr-39 to Lys-58.

The tissue distribution in keratinocytes, combined with the homology to the bovine acrosin inhibitors IIa and IIb indicates that polynucleotides and polypeptides

corresponding to this gene are useful for the acceleration of wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:120 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 482 of SEQ ID NO:120, b is an integer of 15 to 496, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:120, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 111

This gene is expressed primarily in fetal liver/spleen, T cells, and to a lesser extent, in bone marrow and primary dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders; immune dysfunction; lymphomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:270 as residues: Glu-28 to His-34.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of hematopoietic disorders. This gene product is primarily expressed in hematopoietic cells

and tissues, suggesting that it plays a role in the survival, proliferation, and/or differentiation of hematopoietic lineages. This is particularly supported by the expression of this gene product in fetal liver and bone marrow, the two primary sites of definitive hematopoiesis. Expression of this gene product in T cells and primary
5 dendritic cells also strongly indicates a role for this protein in immune function and immune surveillance. Furthermore, since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
10 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:121 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
Accordingly, preferably excluded from the present invention are one or more
15 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1160 of SEQ ID NO:121, b is an integer of 15 to 1174, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:121, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 112

The gene encoding the disclosed cDNA is thought to reside on chromosome 14.
Accordingly, polynucleotides related to this invention are useful as a marker in linkage
25 analysis for chromosome 14.

This gene is expressed primarily in fetal liver, spleen, and to a lesser extent in melanocyte.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
30 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, integumentary, or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of fetal and embryonic tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, developmental, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:271 as residues: Met-1 to Met-7, Gln-43 to Glu-50, Thr-89 to Thr-95.

The tissue distribution in fetal liver and spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of developmental hematopoietic disorders. Additionally, the tissue distribution indicates 10 that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of hematopoietic disorders. This gene product is primarily expressed in hematopoietic cells and tissues, suggesting that it plays a role in the survival, proliferation, and/or differentiation of hematopoietic lineages. This is particularly supported by the expression of this gene product in fetal liver, which is a 15 primary sites of definitive hematopoiesis, and strongly suggesting a role for this protein in immune function and immune surveillance.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 20 ID NO:122 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more 25 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1032 of SEQ ID NO:122, b is an integer of 15 to 1046, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:122, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 113

30

When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates T-cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are 35 involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS

element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in B cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B cell lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:272 as residues: Gln-23 to Asn-31, Tyr-42 to Ser-58.

The tissue distribution in B-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of lymphomas, particularly B cell lymphomas. Furthermore, expression of this gene product in B-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Additionally, the biological activity data supports the notion that the

translational product of this gene activates specific immune cells, and therefore may play a role in the initiation of immune system activity.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:123 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1146 of SEQ ID NO:123, b is an integer of 15 to 1160, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:123, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 114

This gene is expressed primarily in neutrophils: IL-1 and LPS induced.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of certain immune disorders, especially those involving neutrophils. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a

usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for
5 immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies
10 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:124 and may have been publicly available prior to conception of the present
15 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 879 of SEQ ID NO:124, b is an integer of 15
20 to 893, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:124, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 115

25

One embodiment of this gene comprises polypeptides of the following amino acid sequence: DIMPASVIFLICEGVLYGVQG (SEQ ID NO:476). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in placenta.
30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, placental insufficiency; developmental abnormalities; aberrant angiogenesis; abnormal development and/or maintenance of the placenta. Similarly,
35 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the placenta and, more

generally, the vasculature and/or endothelium, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developing, placental, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:125 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1035 of SEQ ID NO:125, b is an integer of 15 to 1049, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:125, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 116

35

This gene is expressed primarily in keratinocytes, as well as in synovial hypoxia and T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, integumentary, immune, or skeletal disorders, particularly wound 5 healing and rheumatoid conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the integumentary system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. 10 skin, connective tissues, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:275 as residues: Thr-42 to Pro-53, Val-78 to Glu-86, Glu-103 to Met-112, Ala-124 to Gly-131.

The tissue distribution in keratinocytes indicates that polynucleotides and 20 polypeptides corresponding to this gene are useful for the treatment of integumentary disorders, particularly with regard to wound healing. Furthermore, the tissue distribution also indicates that the translation product of this gene is useful for the treatment and/or detection of disorders of the connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and 25 dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above 30 listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:126 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the 35 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 1612 of SEQ ID NO:126, b is an integer of 15 to 1626, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:126, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 117

This gene is expressed primarily in hepatoma and testes tumor, and to a lesser extent, in brain.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hepatic, neural, or reproductive disorders, particularly metastatic liver cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful 15 in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cancer and metabolic systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. liver, brain, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, seminal fluid, 20 amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in hepatic tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some types of cancer including hepatoma, testes tumor and related metastases. Furthermore, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are 30 attributable to the differentiation of hepatocyte progenitor cells). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

35 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:127 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1163 of SEQ ID NO:127, b is an integer of 15 to 1177, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:127, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 118

This gene is expressed primarily in CD34 positive cells, and to a lesser extent, in pancreatic tumor and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, endocrine, or immune disorders, particularly pancreatic cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the tumor, immune and metabolic systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, liver, spleen, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, bile, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level. i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in pancreatic and CD34 positive cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some types of cancer, especially those involving CD34 cells and pancreatic cancer. Furthermore, expression of this gene product in both CD34 positive cells and spleen indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for

immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, z, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the 5 differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 10 ID NO:128 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 15 a-b, where a is any integer between 1 to 1262 of SEQ ID NO:128, b is an integer of 15 to 1276, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:128, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 119

This gene is expressed primarily in osteoclastoma, fetal liver/spleen, and to a lesser extent, in primary dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as 25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, osteoclastoma; hematopoietic disorders; lymphomas; impaired immunity; immune disorders; inflammation, in addition to integumentary disorders. Similarly, 30 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and bone, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, bone, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, amniotic fluid, plasma, urine, 35 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the

expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:278 as residues: Thr-23 to Pro-29, Thr-68 to Pro-76.

5 The tissue distribution in dendritic cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of bone and hematopoietic disorders. Elevated levels of expression of this gene product in osteoclastoma indicates that it may play a role in the survival, proliferation, and/or growth of osteoclasts. Therefore, it may be useful in influencing bone mass in such
10 conditions as osteoporosis. More generally, as evidenced by expression in fetal liver/spleen, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in augmentation of the numbers of stem cells and committed progenitors. Expression of this gene product in primary dendritic cells also indicates that it may play a role in mediating responses to
15 infection and controlling immunological responses, such as those that occur during immune surveillance.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:129 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1320 of SEQ ID NO:129, b is an integer of 15 to 1334, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:129, and where b is greater than or equal to a + 14.
20
25

FEATURES OF PROTEIN ENCODED BY GENE NO: 120

30 When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 assay. Thus, it is likely that this gene activates fibroblast cells through a signal transduction pathway. Early growth response 1 (EGR1) is a promoter associated with certain genes that induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.
35

This gene is expressed primarily in hemangiopericytoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, soft tissue cancers, such as hemangiopericytoma, in addition to other 5 proliferative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. circulatory system, and 10 cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:279 as residues: Pro-49 to Thr-64.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of hemangiopericytoma. Furthermore, the biological activity data demonstrates that the 20 translation product of this gene activates fibroblast cells. Fibroblast cells have the ability to undergo vascularization, and thus the translation product of this gene may be involved in disorders of the vascular tissue, such as hemangiopericytoma.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 25 ID NO:130 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 30 a-b, where a is any integer between 1 to 518 of SEQ ID NO:130, b is an integer of 15 to 532, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:130, and where b is greater than or equal to a + 14.

35 **FEATURES OF PROTEIN ENCODED BY GENE NO: 121**

This gene is expressed primarily in kidney cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal or urogenital disorders, particularly nephritis. Similarly, 5 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal system, expression of this gene at significantly higher or lower levels may be detected in certain 10 tissues or cell types (e.g. kidney, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:280 as residues: Pro-33 to Ser-38.

The tissue distribution in kidney cortex indicates that polynucleotides and 20 polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases of the kidney, including nephritis. Furthermore, the tissue distribution in kidney indicates that this gene or gene product could be used in the treatment and/or detection of kidney diseases including renal failure, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, 25 polycystic kidney, and Falconi's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available 30 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:131 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more 35 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 671 of SEQ ID NO:131, b is an integer of 15 to 685, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:131, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 122

5 This gene is expressed primarily in spleen from chronic lymphocytic leukemia. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, such as chronic lymphocytic 10 leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. spleen, cancerous and wounded tissues) 15 or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 The tissue distribution in spleen tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of chronic lymphocytic leukemia. Furthermore, the expression observed predominantly in spleen cells also indicates that the polynucleotides or polypeptides are important in treating and/or detecting hematopoietic disorders, such as graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow 25 fibrosis, and myeloproliferative disease. The polypeptides or polynucleotides are also useful to enhance or protect proliferation, differentiation, and functional activation of hematopoietic progenitor cells (e.g., bone marrow cells), useful in treating cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The polypeptides or polynucleotides are also useful to increase the proliferation of 30 peripheral blood leukocytes, which can be used in the combat of a range of hematopoietic disorders, including immunodeficiency diseases, leukemia, and septicemia.

35 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:132 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 715 of SEQ ID NO:132, b is an integer of 15 to 729, where both a and b correspond to the positions of nucleotide residues shown in 5 SEQ ID NO:132, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 123

10 This gene is expressed primarily in neutrophils, dendritic cells, and CD34 positive cells (Cord Blood).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 15 not limited to, immune, hematopoietic, or developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain 20 tissues or cell types (e.g. immune, developmental, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of some types of immune disorders, especially those involving neutrophils. More generally, as evidenced by expression in CD34 positive cells, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, 30 and may be of use in augmentation of the numbers of stem cells and committed progenitors. Expression of this gene product in primary dendritic cells also indicates that it may play a role in mediating responses to infection and controlling immunological responses, such as those that occur during immune surveillance.

Many polynucleotide sequences, such as EST sequences, are publicly available 35 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:133 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1065 of SEQ ID NO:133, b is an integer of 15 to 1079, where both a and b correspond to the positions of nucleotide residues shown 5 in SEQ ID NO:133, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 124

10

This gene is expressed primarily in adult lung.

15

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, respiratory disorders. Similarly, polypeptides and antibodies directed to 20 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. respiratory, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25

The tissue distribution in lung tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of respiratory disorders, such as asthma, emphysema, and ARDS. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:134 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

35

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1283 of SEQ ID NO:134, b is an integer of 15

to 1297, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:134, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 125

The gene encoding the disclosed cDNA is thought to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

10 This gene is expressed primarily in T-cell lymphoma and fetal liver/spleen. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, developmental, or hematopoietic disorders, particularly 15 lymphomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, developmental, and cancerous 20 and wounded tissues) or bodily fluids (e.g.lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:284 as residues: Gln-25 to Phe-43.

The tissue distribution in T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of T-cell lymphoma. Furthermore, expression of this gene product in fetal liver/spleen indicates a role in the 30 regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, 35 the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma,

immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:135 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 603 of SEQ ID NO:135, b is an integer of 15 to 617, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:135, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 126

20

The translation product of this gene shares sequence homology with C9, a gene of unknown function. The gene encoding the disclosed cDNA is thought to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3. One embodiment of this gene comprises the polypeptides of the following amino acid sequence:

GTAFQHAFSTNDCSRNVYIKKNGFTLHRNPIAQSTIDGARTKIGFSEGRHAWEV
WWEGPLGTVAVIGIATKRAPMQCQGYVALLGSDDQSWGNLVDNNLLHN
VNGSFPQCNNAPKYQIGERIRVILDMEDKTLAFERGYEFLGVAFRGLPKVCLYP
AVSAVYGNTEVTLVYLGKPLDG (SEQ ID NO:477). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in placenta, and to a lesser extent, in apoptotic T-cells, as well as in smooth muscle, testes, and microvascular endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, reproductive, muscular, vascular, and cancerous and wounded tissues) 5 or bodily fluids (e.g. lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in T-cells combined with the homology to the C9 protein 10 indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some immune disorders, especially those involving T-cells. Furthermore, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness 15 in the treatment of cancer (e.g. by boosting immune responses), or male infertility. Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available 20 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:136 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 25 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1297 of SEQ ID NO:136, b is an integer of 15 to 1311, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:136, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 127

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as 35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some immune disorders, especially those involving neutrophils. Furthermore, as evidenced by expression in neutrophils, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in augmentation of the number of stem cells and committed progenitors. Expression of this gene product in neutrophils further indicates that it may play a role in mediating responses to infection and controlling immunological responses, such as those that occur during immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:137 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1081 of SEQ ID NO:137, b is an integer of 15 to 1095, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:137, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 128

This gene is expressed primarily in neutrophils; IL-1 and LPS induced. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, immune or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:287 as residues: Lys-36 to Asp-42.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some immune disorders, especially those involving neutrophils. Furthermore, as evidenced by the expression in neutrophils, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in augmentation of the number of stem cells and committed progenitors. Expression of this gene product in neutrophils further indicates that it may play a role in mediating responses to infection and controlling immunological responses, such as those that occur during immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:138 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 678 of SEQ ID NO:138, b is an integer of 15 to 692, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:138, and where b is greater than or equal to a + 14.

This gene is expressed primarily in neutrophils, IL-1 and LPS induced.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:288 as residues: Pro-32 to Gln-38, Gly-51 to Asp-57.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of certain immune disorders, especially those involving neutrophils. Furthermore, as evidenced by expression in neutrophils, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in augmentation of the number of stem cells and committed progenitors. Expression of this gene product in neutrophils further indicates that it may play a role in mediating responses to infection and controlling immunological responses, such as those that occur during immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:139 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 734 of SEQ ID NO:139, b is an integer of 15

to 748, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:139, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 130

This gene is expressed primarily in neutrophils, IL-1 and LPS induced.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at
15 significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
20 having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:289 as residues: Gly-22 to Ser-28.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of
25 certain immune disorders involving neutrophils. Furthermore, as evidenced by expression in neutrophils, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in augmentation of the number of stem cells and committed progenitors. Expression of this gene product in neutrophils further indicates that it may play a role in mediating responses to infection
30 and controlling immunological responses, such as those that occur during immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
35 ID NO:140 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1118 of SEQ ID NO:140, b is an integer of 15 to 1132, where both a and b correspond to the positions of nucleotide residues shown 5 in SEQ ID NO:140, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 131

10 This gene is expressed primarily in corpus callosum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural disorders, particularly diseases of the brain, such as 15 neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. brain, and 20 cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in neural tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of brain disorders and diseases, including paranoia, schizophrenia, depression, mania, and Alzheimer's disease. Furthermore, elevated expression of this gene product within the corpus callosum of the brain indicates that it may be involved in neuronal survival; 30 synapse formation; conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. Again, it may also be useful in the treatment of such neurodegenerative disorders as schizophrenia; ALS; or Alzheimer's. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

35 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:141 and may have been publicly available prior to conception of the present

invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 5 a-b, where a is any integer between 1 to 1098 of SEQ ID NO:141, b is an integer of 15 to 1112, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:141, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 132

The translation product of this gene shares sequence homology with the putative transposase of the Tigger-1 transposon.

This gene is expressed primarily in atrophic endometrium.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, muscular disorders, particularly muscular atrophy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing 20 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. reproductive, muscular, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid 25 and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in endometrial tissue combine with the homology to a transposase indicates that polynucleotides and polypeptides corresponding to this gene 30 are useful for DNA repair in atrophying tissue, particularly of the endometrium.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 35 ID NO:142 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1070 of SEQ ID NO:142, b is an integer of 15 to 1084, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:142, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 133

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ARAFQHLMVADHSFHRTLIKQPSMIPNATFYHIF (SEQ ID NO:478). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in hemangiopericytoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, soft tissue tumors, particularly hemangiopericytoma, or other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:292 as residues: Ser-39 to Ser-44.

The tissue distribution in hemangiopericytoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of various soft-tissue tumors, in addition to other proliferative disorders which may afflict other tissues or cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:143 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more

- 5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1036 of SEQ ID NO:143, b is an integer of 15 to 1050, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:143, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 134

This gene is expressed primarily in hypothalamus of a schizophrenic patient, and to a lesser extent in spleen.

- 15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural or immune disorders, particularly Schizophrenia or neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these
20 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, immune, hematopoietic, spleen, cancerous and wounded tissues) or bodily
25 fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 30 The tissue distribution in hypothalamus indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of Schizophrenia, as well as other central nervous system and immune system disorders. Furthermore, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease,
35 Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania.

dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, disorders of the endocrine system, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:144 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1099 of SEQ ID NO:144, b is an integer of 15 to 1113, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:144, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 135

The translation product of this gene shares sequence homology with a chicken ring-finger-zinc finger protein, C-RZF, in addition to, the human multiple membrane spanning receptor TRC8 which is thought to serve as a signaling receptor in renal and thyroid carcinomas. (See Genbank Accession No.gil3395787 (AF064801)) The TRC8 locus has been described in a family with classical features of hereditary renal cell carcinoma. The 8q24.1 (locus of TRC8) breakpoint region encodes the 664-aa multiple membrane spanning protein, TRC8, with similarity to the hereditary basal cell carcinoma/segment polarity gene, patched. This similarity involves two regions of patched, the putative sterol-sensing domain and the second extracellular loop that participates in the binding of sonic hedgehog. In the 3;8 translocation, TRC8 is fused to FHIT (fragile histidine triad gene) and is disrupted within the sterol-sensing domain. In

contrast, the FHIT coding region is maintained and expressed. In a series of sporadic renal carcinomas, an acquired TRC8 mutation was identified. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

ARALPEIKGSRLQEINDVCAICYHEFTTSARITPCNHYFHALCLRKWLYIQDTC

5 MCHQKVYIEDDIKDN

SNVSNNNGFIPPNETPEEA VREAAAESDRELNE DDSTDCDDDVQRERNGVIQHT

GAAAGRI (SEQ ID NO:479), FSTQAQQLEEFNDDTD (SEQ ID NO:480), RLQE

INDVCAICYHEFTTSARI (SEQ ID NO:481), LYIQDTCPMCHQKVYIEDDI (SEQ

ID NO:482), VSNNNGFIPPNETPEEA VREA (SEQ ID NO:483), and/or DDSTD CD

10 DDVQRERNGVIQHTGAAAG (SEQ ID NO:484). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 8. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 8.

This gene is expressed primarily in human embryonic tissues.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental abnormalities, particularly congenital defects or proliferative conditions. Similarly, polypeptides and antibodies directed to these 20 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the embryonic tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developmental, renal, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic 25 fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in embryonic tissue, combined with the homology to ring finger-zinc finger protein and the human TRC8 receptor indicates that polynucleotides 30 and polypeptides corresponding to this gene are useful for diagnosis and treatment of abnormalities of the embryonic tissues, in particular proliferative disorders. In addition, polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, detection, and/or treatment of developmental disorders. The relatively 35 specific expression of this gene product during embryogenesis indicates that it may be a key player in the proliferation, maintenance, and/or differentiation of various cell types during development. It may also act as a morphogen to control cell and tissue type

- specification. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. Moreover, this protein may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.
- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:145 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 671 of SEQ ID NO:145, b is an integer of 15 to 685, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:145, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 136

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: VAGITGAHHHAQLIFVLLVEMGFHHV GQAGLKLLTSDN PRTSASQSAGITGMSXGRRITCGQEFKTAVSYNCTTALQPDRAKLCFLFKKKK KISIQ RTLPGIKRVIYNYERVDSSKGHNSQVQWAHA CNPSTLGGRRGGQIV (SEQ ID NO:485), AGITGAHHHAQLIFVLLVEMGF (SEQ ID NO:486), RVIYN YERVDSSKGHNSQVQWAHACNP (SEQ ID NO:487). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in microvascular endothelial cells. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular or endothelial disorders, such as the following: arteriosclerosis, tumorigenesis, stroke, embolism, aneurysm, microvascular disease, and various cardiovascular disorders. Similarly, polypeptides and antibodies directed to these

- polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. vascular, endothelial, 5 cardiovascular, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 10 The tissue distribution in microvascular endothelial tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of vascular disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.
- 15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:146 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 20 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1024 of SEQ ID NO:146, b is an integer of 15 to 1038, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:146, and where b is greater than or equal to a + 14.
- 25

FEATURES OF PROTEIN ENCODED BY GENE NO: 137

- The gene encoding the disclosed cDNA is believed to reside on chromosome 2.
- 30 Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2.

This gene is expressed primarily in fetal tissues, most notably fetal cochlea and fetal lung, and to a lesser extent, in rhabdomyosarcoma and healing groin wound tissue.

- 35 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, embryological/developmental abnormalities; hearing defects; respiratory diseases; rhabdomyosarcoma; general cancers and other proliferative conditions; fibrosis; wound healing. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification 5 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the embryo/fetus or of striated muscle cells, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. developmental, pulmonary, auditory, muscle, fibroid, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal 10 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder

The tissue distribution in fetal tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for diseases involving abnormal 15 cellular proliferation, such as cancer. Expression of this gene product in rapidly proliferating cells, such as those found in the embryo; in rhabdomyosarcomas; and in wound healing tissue, indicates that this gene may play a role in controlling or promoting cell proliferation. Alternately, expression of this gene in fetal tissues indicates that it may play a role in cellular development and differentiation, particularly 20 of the auditory system as well as the lungs. Thus, this gene product may be useful in the treatment and/or diagnosis of hearing defects, as well as respiratory disorders. Expression of this gene product in rhabdomyosarcoma indicates that it may play a role 25 in the progression of such cancers, and may also be involved in metastasis and/or angiogenesis. Additionally, expression in wound healing tissues again indicates a role in the proliferation of connective tissue types involved in wound healing, as well as in the fibrosis and scarring that accompanies the wound healing process. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available 30 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:147 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 35 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 837 of SEQ ID NO:147, b is an integer of 15

to 851, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:147, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 138

The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

10 This gene is expressed primarily in adult brain, and to a lesser extent, in cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 15 not limited to, disorders and diseases of the brain, particularly neurodegenerative and behavior conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher 20 or lower levels may be detected in certain tissues or cell types (e.g. neural, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:297 as residues: Pro-25 to Ser-30, Thr-36 to Ser-47.

The tissue distribution in neural tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders and diseases of the brain, particularly paranoia, Alzheimer's, depression, 30 schizophrenia, and mania. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal 35 cord injuries, ischemia and infarction, aneurysms, hemorrhages, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance,

and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product
5 may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:148 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

15 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 600 of SEQ ID NO:148, b is an integer of 15 to 614, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:148, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 139

This gene is expressed primarily in cerebellum.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural disorders, particularly neurodegenerative disorders, such as Alzheimers. Similarly, polypeptides and antibodies directed to these polypeptides are
30 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.neural, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in cerebellum indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of brain diseases and disorders. Specifically, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:149 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1186 of SEQ ID NO:149, b is an integer of 15 to 1200, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:149, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 140

This gene is expressed primarily in brain tissue of a patient with Alzheimer's disease, and to a lesser extent, in human adipose tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural or adipose-related disorders, particularly neurodegenerative disorders, such as Alzheimer's disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, metabolic, adipose, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in neural and adipose tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of Alzheimer's disease and other nervous system disorders. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. More specifically, the protein product of this gene may show utility in the treatment, diagnosis, and/or prevention of neural disorders which occur secondary to aberrations in fatty-acid metabolism, such as improper development of the myelin sheath of nerve cells, for example. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:150 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more

- 5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 669 of SEQ ID NO:150, b is an integer of 15 to 683, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:150, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 141

This gene is expressed primarily in T cells.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly T cell leukemia, immunodeficiencies, and inflammatory conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:300 as residues: Asn-62 to Leu-68.

- 30 The tissue distribution T-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of T cell leukemia and other disorders of the immune system. Moreover, this gene product may play a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene

product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, 5 hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene 10 product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available 15 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:151 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 20 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 813 of SEQ ID NO:151, b is an integer of 15 to 827, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:151, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 142

The gene encoding the disclosed cDNA is believed to reside on chromosome 8. Accordingly, polynucleotides related to this invention are useful as a marker in linkage 30 analysis for chromosome 8.

This gene is expressed primarily in the frontal lobe of the brain, and to a lesser extent, in synovial fluid and embryos.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 35 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental or neural disorders, particularly neurodegenerative, behavioral, and congenital abnormalities of the brain. Similarly, polypeptides and

antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell

5 types (e.g. neural, developmental, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:301 as residues: Gln-24 to Lys-31.

The tissue distribution in brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of abnormalities of the brain. Moreover, polynucleotides and polypeptides corresponding 15 to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, 20 hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in 25 synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the skeletal or cardiovascular system. Protein, as well as, antibodies directed against the protein may 30 show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:152 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 821 of SEQ ID NO:152, b is an integer of 15 to 835, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:152, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 143

The gene encoding the disclosed cDNA is believed to reside on chromosome 10 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in osteoblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal disorders, such as osteoporosis, and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, 20 expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g.skeletal, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an 25 individual not having the disorder.

The tissue distribution in osteoblasts indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of osteoporosis and other bone degenerative diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above 30 listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 35 ID NO:153 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 544 of SEQ ID NO:153, b is an integer of 15 to 558, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:153, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 144

This gene is expressed primarily in CD34 positive cells (cord blood) and placenta.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and immune disorders, particularly proliferative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are 15 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, reproductive, developmental, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, 20 serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in cord blood and placental tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of certain immune disorders, especially those involving CD34 cells. Expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, 30 developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

35 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:154 and may have been publicly available prior to conception of the present

invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 5 a-b, where a is any integer between 1 to 1187 of SEQ ID NO:154, b is an integer of 15 to 1201, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:154, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 145

This gene is expressed primarily in frontal cortex of the brain. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural or spinal cord disorders, such as neurodegenerative conditions and other abnormalities of the brain. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above 20 tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression 25 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:304 as residues: Pro-21 to Ser-27.

The tissue distribution in frontal cortex tissue indicates that polynucleotides and 30 polypeptides corresponding to this gene are useful for diagnosis and treatment of the abnormalities of the brain. Specifically, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, 35 encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive

disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in
5 synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a
10 tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:155 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
15 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1012 of SEQ ID NO:155, b is an integer of 15 to 1026, where both a and b correspond to the positions of nucleotide residues shown
20 in SEQ ID NO:155, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 146

25 The gene encoding the disclosed cDNA is believed to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2.

This gene is expressed primarily in adrenal gland tumor, breast tissue, and to a lesser extent in adipose tissue.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine or reproductive disorders, such as adrenal gland tumor; breast cancer; metabolic disorders. Similarly, polypeptides and antibodies directed to these
35 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adrenal glands and breast, expression of this gene at significantly

higher or lower levels may be detected in certain tissues or cell types (e.g.reproductive, metabolic, endocrine, breast, adrenal gland, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, breast milk, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, 5 relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:305 as residues: Arg-44 to Lys-49, Asp-60 to Phe-66.

The tissue distribution in adrenal gland and breast tissues indicates that 10 polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders involving the adrenal gland. Expression of this gene product in adrenal gland tumor indicates that it may play a role in the proliferation of cells of the adrenal gland, or potentially in the proliferation of cells in general. In such an event, it may play a role in determining the course and severity of cancer. 15 Alternatively, it may play a role in the normal function of adrenal glands, such as in the production of corticosteroids, androgens, or epinephrines. Thus it may play a role in general homeostasis, as well as in disorders involving the androgen hormones. Expression of this gene product in breast and adipose tissues also indicates that it may play a role in breast cancer, or in supplying vital nutrients to the infant during lactation. 20 Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:156 and may have been publicly available prior to conception of the present 25 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 890 of SEQ ID NO:156, b is an integer of 15 30 to 904, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:156, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 147

35

This gene is expressed primarily in LNCAP, and untreated spleen; metastatic melanoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, integumentary disorders, such as metastatic melanoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and cancer metabolic systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. immune, hematopoietic, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:306 as residues: His-47 to Thr-53.

The tissue distribution in spleen and integumentary tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of some types of cancer, especially metastatic melanoma. The protein product of this gene is useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type

II, metaphyseal chondrodysplasia type Schmid). Alternatively, this gene is useful for the treatment and diagnosis of hematopoetic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone 5 marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various 10 blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 15 ID NO:157 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 20 a-b, where a is any integer between 1 to 902 of SEQ ID NO:157, b is an integer of 15 to 916, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:157, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 148

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: AGAEVVVMLFLTPSS HHQHECVRRAFECGDCHILLDNNV 30 LGVDCHGAGERAVHLEDHFVHIDTISLLDEALEYSALIAGHPKSD LPPGLSRC RPWEHHWPISYTG (SEQ ID NO:488), TISYLCNNVSYMQLQKLVGKSMIFLP YSLPIHLPGNHRLLLPRVGMRLRGCCFSPYIITDFKWC (SEQ ID NO:489), EMGQWCSQGLHLDSPGGKSDFGCPAINAEYSRASSKSRLMVSMWTKWSSRC TALSPAP (SEQ ID NO:490), RAFECGDCHILLDNNVLGVVDCHGAG (SEQ ID NO:491), and/or LVGKSMIFLPYSLPIHLPGNHRL (SEQ ID NO:492).
35 Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 1.

Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in ovary, and to a lesser extent in meninges, the adrenal gland, and the cerebellum.

- 5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, neural, and endocrine disorders, such as ovarian and brain cancers, neurodeficiency disorders, and infertility. Similarly, polypeptides and
- 10 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive and endocrine systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g. neural, reproductive, ovarian, and cancerous and
- 15 wounded tissues) or bodily fluids (e.g. lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 20 The tissue distribution in ovarian and endocrine tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of ovarian cancer and other endocrine disorders. Alternatively, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or
- 25 inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning
- 30 disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or
- 35 survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well

as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 5 ID NO:158 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 10 a-b, where a is any integer between 1 to 907 of SEQ ID NO:158, b is an integer of 15 to 921, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:158, and where b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID Total NO: NT Seq. X	5' NT of Clone Seq. Start Codon	3' NT First AA of Signal Pep Y	5' NT First AA of ID NO: Sig Pep	First AA of AA Sig Portion	Last AA of AA Secreted Portion	Last AA of AA Secreted Portion ORF
1	HNGEU17	209299 09/25/97	Uni-ZAP XR	11	826	1	826	277	277	160
2	HNGDJ72	209299 09/25/97	Uni-ZAP XR	12	524	1	524	185	185	161
3	HNGEO29	209299 09/25/97	Uni-ZAP XR	13	491	1	491	98	98	162
4	HNHDL95	209299 09/25/97	Uni-ZAP XR	14	403	1	403	121	121	163
5	HAGDS35	209299 09/25/97	Uni-ZAP XR	15	813	1	813	52	52	164
6	HNGEQ48	209299 09/25/97	Uni-ZAP XR	16	264	1	264	10	10	165
7	HNGDG40	209299 09/25/97	Uni-ZAP XR	17	520	1	520	13	13	166

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	5' NT of Clone Seq. of Seq.	3' NT of Clone Seq. Start Codon	AA of AA of Signal NO: Y	First SEQ ID of Pep	Last AA of Sig	First AA of Sig	Last AA of Pep	First AA of Secreted Portion	Last AA of ORF
8	HNGEN81	209299	Uni-ZAP XR 09/25/97	18	993	1	993	380	167	1	25	26
9	H2MAC30	209299	pBluescript SK- 09/25/97	19	459	1	459	157	168	1	28	29
10	hhfb16	209299	Uni-ZAP XR 09/25/97	20	555	1	555	344	344	1	23	24
11	HPFCL43	209299	Uni-ZAP XR 09/25/97	21	665	1	665	21	21	170	1	17
12	HSATR82	209299	Uni-ZAP XR 09/25/97	22	777	1	777	74	74	171	1	15
13	H6EDF66	209299	Uni-ZAP XR 09/25/97	23	540	1	540	146	146	172	1	27
14	HNHIC21	209299	Uni-ZAP XR 09/25/97	24	484	1	484	65	65	173	1	16
15	HOVCA92	209299	pSportI 09/25/97	25	707	1	488	181	181	174	1	20
												62

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	NT of 3' NT of 5' NT First AA of AA of ID of Signal NO: Sig Y	5' NT of 3' NT First AA of AA of ID of Signal NO: Sig Y	First AA of AA of ID of Signal NO: Sig Y	Last AA of AA of ID of Signal NO: Sig Y	Last AA of AA of ID of Signal NO: Sig Y
				Start Seq.	Codon Pep	Pep	Pep	Portion ORF
16	HNHDW38	209299 09/25/97	Uni-ZAP XR	26	793	1	793	66
17	HSDL30	209299 09/25/97	Uni-ZAP XR	27	638	1	638	26
18	HATDB65	209299 09/25/97	Uni-ZAP XR	28	528	14	528	110
19	HPMSM14	209299 09/25/97	pBluescript	29	919	1	919	119
20	HTTEA24	209299 09/25/97	Uni-ZAP XR	30	864	1	864	133
21	HAGDS20	209299 09/25/97	Uni-ZAP XR	31	919	1	919	11
22	HSDJM30	209299 09/25/97	Uni-ZAP XR	32	956	1	956	70
23	HNHEE88	209299 09/25/97	Uni-ZAP XR	33	566	1	566	87

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	5' NT of Total Clone Seq. Start Seq.	3' NT of AA SEQ ID of AA Signal NO: Pep Y	Last AA of AA Secreted Portion ORF	
24	HSLFD55	209346	Uni-ZAP XR 10/09/97	34 1564 1	1035 129	105 183 1	21 22 43
25	HSAXJ29	209299	Uni-ZAP XR 09/25/97	35 1035 1	1035 129	129 184 1	19 20 57
26	HSFAM39	209299	Uni-ZAP XR 09/25/97	36 620 1	620 117	117 185 1	23 24 68
27	HTODO72	209299	Uni-ZAP XR 09/25/97	37 973 1	973 183	183 186 1	16 17 24
28	HADDZ85	209299	pSport1 09/25/97	38 838 1	838 270	270 187 1	36 37 57
29	HDPCM26	209300	pCMV Sport 3.0 09/25/97	39 607 1	607 174	174 188 1	19 20 66
30	HSZAA13	209300	Uni-ZAP XR 09/25/97	40 882 1	855 147	147 189 1	19 20 88
31	HTDBP04	209300	pCMV Sport 2.0 09/25/97	41 959 1	959 65	65 190 1	15 16 220

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: NT Seq. X	5' NT of Total NT Seq. NT Seq.	3' NT of Clone Seq.	5' NT of AA of ID of AA of Seq. Start Codon	5' NT of AA of ID of Signal Pep.	First AA of NO: Y	Last AA of Sig Pep.	First AA of Sig Pep.	Last AA of Portion ORF
32	HHGQ54	209300 09/25/97	Lambda ZAP II	42 875	1	875	62	191	1	15	16 51
33	HSNAB12	209300 09/25/97	Uni-ZAP XR	43 630	1	630	151	192	1	27	28 71
34	HBJID05	209300 09/25/97	Uni-ZAP XR	44 571	1	571	137	193	1	20	21 111
35	HSNBM49	209300 09/25/97	Uni-ZAP XR	45 930	1	930	27	194	1	21	22 60
36	HJMBF77	209300 09/25/97	pCMV Sport 3.0	.46 437	1	432	60	195	1	24	25 126
37	HJMBM38	209300 09/25/97	pCMV Sport 3.0	.47 1024	316	1023	387	196	1	15	16 112
38	HHGCL33	209300 09/25/97	Lambda ZAP II	.48 463	1	463	74	197	1	20	21 65
39	HCEWE20	209300 09/25/97	Uni-ZAP XR	.49 885	13	885	166	198	1	18	19 51

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NT Seq. X Seq.	5' NT of Total Clone Seq. Start Seq.	3' NT of Clone Seq. Signal Codon Pep Y	5' NT of AA First SEQ ID of AA of Signal NO: Pep Y	First AA of AA Sig Pep Y	Last AA of AA Sig Portion	First AA of AA Sig Portion	Last AA of AA Sig Portion	Last ORF
40	HCUHL13	209300 09/25/97	ZAP Express	50 847 1	847 84	84 199	1 20	21	21	21	58
41	HBJHO68	209300 09/25/97	Uni-ZAP XR	51 580 1	580 34	34 200	1 1	24	24	25	51
42	HCWDV84	209300 09/25/97	ZAP Express	52 598 1	598 47	47 201	1 1	25	25	26	80
43	HBXFC78	209300 09/25/97	ZAP Express	53 571 1	567 184	184 202	1 1	14	14	15	69
44	HE2F145	209300 09/25/97	Uni-ZAP XR	54 1247 212	1082 273	273 203	1 1	38	38	39	45
45	HEOMG13	209300 09/25/97	pSport1	55 848	182 848	247 247	1 1	27	27	28	52
46	HFAMH77	209300 09/25/97	Uni-ZAP XR	56 669 96	669 240	240 205	1 1	33	33	34	61
47	HSVCF20	209300 09/25/97	Uni-ZAP XR	57 680 1	680 43	43 206	1 1	25	25	26	43

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NT NO: NT Seq. X	5' NT of Clone Seq. Start Seq.	3' NT of AA First SEQ ID NO: Signal Pep Y	5' NT of AA First AA of ID Sig Pep	Last AA of Pep	First AA of Sig Portion	Last AA of Secreted Portion	First AA of ORF
48	HISAG02	209300 09/25/97	pSport I	58	524 1	524 18	18 207	1 27	28	40
49	HCDAF84	209300 09/25/97	Uni-ZAP XR	59	427 1	427 168	168 208	1 18	19	56
50	HHAAC17	209300 09/25/97	Uni-ZAP XR	60	1263 1	1263 227	227 209	1 19	20	125
51	HSNMC45	209300 09/25/97	Uni-ZAP XR	61	720 1	720 232	232 210	1 19	20	25
52	HEQAG39	209300 09/25/97	pCMV Sport	62	589 69	589 93	93 211	1 19	20	47
53	HKACH44	209300 09/25/97	pCMV Sport	63	686 1	686 375	375 212	1 25	26	44
54	HBNBG49	209300 09/25/97	Uni-ZAP XR	64	452 1	452 40	40 213	1 34	35	51
55	HE2EN04	209300 09/25/97	Uni-ZAP XR	65	370 1	370 57	57 214	1 16	17	50

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	5' NT of Clone Seq. Start Seq. AA	5' NT of AA First SEQ AA	5' NT of AA First SEQ AA	First AA of ID of Signal NO: Sig Pep Y	Last AA of ID of Signal NO: Sig Pep Y	Secreted Portion	Last AA of ORF
56	HSVAA10	209300 09/25/97	Uni-ZAP XR 66 987	1 987	38	38	215 1	16	17	209
57	HFPBA88	209300 09/25/97	Uni-ZAP XR 67 1018	284 1018	33	33	216 1	38	39	195
57	HFPBA88	209300 09/25/97	Uni-ZAP XR 159 804	70 804	98	98	308 1	41	42	102
58	HFTBM50	209300 09/25/97	Uni-ZAP XR 68 762	1 740	158	158	217 1	20	21	34
59	HHEBW54	209300 09/25/97	pCMV Sport 69 630	1 630	97	97	218 1	37	38	71
60	HFEBH21	209300 09/25/97	Uni-ZAP XR 70 940	1 940	21	21	219 1	30	31	52
61	HFTDZ36	209300 09/25/97	Uni-ZAP XR 71 1103	231 1103	547	547	220 1	22	23	68
62	HGLAW96	209300 09/25/97	Uni-ZAP XR 72 899	246 899	308	308	221 1	24	25	68

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	5' NT of Total Clone Seq. Seq.	3' NT of AA of ID of Signal NO: Pep Y	5' NT of AA First SEQ AA of ID of AA Sig Pep Y	Last AA of AA First SEQ AA of ID of AA Sig Pep Y	Last AA of AA First SEQ AA of ID of AA Sig Secreted Portion ORF
63	HKAFK41	209300 09/25/97	pCMV Sport 2.0	73 549	1 549	590 243	232 243	222 222
64	HOSEG51	209324 10/02/97	Uni-ZAP XR	74 590	48 590	232 232	223 223	31 1
65	HTEJT39	209324 10/02/97	Uni-ZAP XR	75 1056	1 1056	146 146	224 224	32 1
66	HPTRH45	209324 10/02/97	pBluescript	76 930	1 930	92 92	225 225	26 1
67	HDHMA72	209324 10/02/97	pCMV Sport 2.0	77 4463	216 2158	287 287	226 226	27 1
68	HNTBL27	209324 10/02/97	pCMV Sport 3.0	78 791	71 791	100 100	227 227	37 1
69	HCFMX35	209324 10/02/97	pSportI	79 1292	1 1292	160 160	228 228	21 1
70	HMSFS21	209324 10/02/97	Uni-ZAP XR	80 1283	1 1283	28 28	229 229	17 1
								18 37

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NT NO: NT Seq. X	5' NT of Total Clone Seq. Seq.	3' NT of Clone Seq. Seq.	5' NT of AA First SEQ ID of AA of Signal NO: Sig Pep Y	Last AA of AA of Sig Pep Y	First AA of AA of Sig Secreted Portion	Last AA of AA of Sig Secreted Portion	First AA of AA of Sig Secreted Portion	Last AA of AA of Sig Secreted Portion			
								Start Seq.	Clone Seq.	Start Codon	AA First SEQ ID of AA of Signal NO: Sig Pep Y			
87	HMDAP35	209324	Uni-ZAP XR	97	628	1	628	70	70	246	1	21	22	50
88	HTOJK60	209324	Uni-ZAP XR	98	904	1	904	217	217	247	1	19	20	32
89	HWBCN75	209324	pCMV Sport 3.0	99	576	1	576	182	184	248	1	34	35	48
90	HROAH06	209324	Uni-ZAP XR	100	713	1	713	29	29	249	1	43	44	115
91	HSAXA83	209324	Uni-ZAP XR	101	649	1	649	92	92	250	1	22	23	74
92	HSDJE10	209324	Uni-ZAP XR	102	697	1	697	157	157	251	1	21	22	62
93	HBAMA40	209324	pSport1	103	1288	1	1288	95	95	252	1	31	32	72
94	HBAMB34	209324	pSport1	104	1027	1	1027	87	87	253	1	35	36	48

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NT Seq. X Seq.	5' NT of 3' NT of 5' NT	Clone Seq. Start Codon Pep Y	5' NT of AA SEQ ID Signal NO: Sig Pep	First AA of ID of Y Sig Pep	Last AA of Portion ORF
95	HCWKC15	209324 10/02/97	ZAP Express 105 710 1	710 37	37	254 1	18	19 40
96	HDTDM65	209324 10/02/97	pCMV Sport 106 530 1	530 159	159	255 1	40	41 53
97	HMMBF71	209324 10/02/97	pSport 1 107 392	392 153	153	256 1	24	25 40
98	HPBDH4I	209324 10/02/97	pBluescript 108 991	288 991	373	373 257 1	15	16 41
99	HPBEN24	209324 10/02/97	pBluescript SK- 109 912	363 912	541	541 258 1	20	21 52
100	HCUIM65	209324 10/02/97	ZAP Express 110 875	331 736	557	557 259 1	27	28 47
101	HKNAA95	209324 10/02/97	pBluescript SK- 111 459 1	459 114	114	114 260 1	28	29 52
102	HKIYH57	209324 10/02/97	pBluescript 112 609	156 609	336	336 261 1	23	24 54

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	5' NT of 3' NT of Total Clone Seq. Start Seq. AA of Signal NO: Y	5' NT of AA First SEQ ID of Pep	Last AA of AA First SEQ ID of Sig	First AA of AA First SEQ ID of Sig	Last AA of AA First SEQ ID of Sig	Last AA of AA First SEQ ID of Sig
									Secreted Portion ORF
103	HBIBW67	209324	Uni-ZAP XR	113	1404	1	1404	685	685
		10/02/97						262	1
104	HCFCU88	209324	pSport1	114	853	1	853	326	326
		10/02/97						1	14
105	HBJMG49	209324	Uni-ZAP XR	115	845	1	804	53	53
		10/02/97						264	1
106	H6EDC19	209324	Uni-ZAP XR	116	760	324	760	389	389
		10/02/97						1	17
107	HSKHZ81	209346	pBluescript	117	988	1	967	57	57
		10/09/97						266	1
108	HBJJFX78	209346	Uni-ZAP XR	118	1947	1	1947	34	34
		10/09/97						267	1
109	HEMFS60	209346	Uni-ZAP XR	119	1448	63	1448	111	111
		10/09/97						268	1
110	HKACB56	209346	pCMV Sport	120	496	1	496	27	27
		10/09/97	2.0					269	1
								23	24
									80

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X Seq.	5' NT of Total Clone Seq. Start Codon Pep Y	First AA SEQ ID of AA of Signal NO: Sig Pep	Last AA of ID of Y Sig	Last AA of Portion	Last AA of Secreted ORF
111	HTXJX80	209346 10/09/97	Uni-ZAP XR	121 1174 16 880	206	206	270 1 26	27 68
112	HAFBD61	209346 10/09/97	pBluescript SK-	122 1046 1 1046	210	210	271 1 22	23 130
113	HBJJU28	209346 10/09/97	Uni-ZAP XR	123 1160 1 1160	133	133	272 1 18	19 84
114	HNHEI47	209346 10/09/97	Uni-ZAP XR	124 893 1 893	192	192	273 1 18	19 78
115	HPMFY74	209346 10/09/97	Uni-ZAP XR	125 1049 1 1049	91	91	274 1 40	41 53
116	HKACD58	209346 10/09/97	pCMV Sport 2.0	126 1626 1 1626	35	35	275 1 25	26 154
117	HLDBB60	209346 10/09/97	pCMV Sport 3.0	127 1177 1 1177	283	283	276 1 20	21 128
118	HLYAP91	209346 10/09/97	pSport1	128 1276 1 1276	280	280	277 1 29	30 83

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	NT SEQ ID Total NO: NT Seq. X	5' NT of Clone Seq. Start Codon	5' NT of Clone Seq. Start Codon	5' NT of AA First SEQ AA ID of Signal NO: Pep Y	AA First SEQ AA ID of Sig Pep Y	Last AA of AA First SEQ AA ID of Sig Pep Y	First AA of AA First SEQ AA ID of Sig Pep Y	Last AA of AA First SEQ AA ID of Sig Pep Y	First AA of AA First SEQ AA ID of Sig Pep Y	Last AA of AA First SEQ AA ID of Sig Pep Y	First AA of AA First SEQ AA ID of Sig Pep Y	Last AA of AA First SEQ AA ID of Sig Pep Y
135	HE8ER60	209346 10/09/97	Uni-ZAP XR	145 685	1	685	48	294	1	32	33	33	74		
136	HMEJQ66	209346 10/09/97	Lambda ZAP II	146 1038	1	1038	80	80	295	1	24	25	25	50	
137	HRDAD66	209346 10/09/97	Uni-ZAP XR	147 851	99	851	269	269	296	1	33	34	34	44	
138	HCMST14	209346 10/09/97	Uni-ZAP XR	148 614	1	614	136	136	297	1	24	25	25	47	
139	HCEBA03	209346 10/09/97	Uni-ZAP XR	149 1200	1	1200	76	76	298	1	21	22	22	54	
140	HFAAH18	209346 10/09/97	Uni-ZAP XR	150 683	79	683	304	304	299	1	21	22	22	29	
141	HJAAM10	209346 10/09/97	pBluescript SK-	151 827	135	827	320	320	300	1	35	36	36	72	
142	HFIBV09	209346 10/09/97	pSport1	152 835	129	835	370	370	301	1	17	18	17	36	

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID Total NO: NT Seq. X	5' NT of Clone Seq. Start Seq. of AA of Signal NO: Pep Y	5' NT of AA SEQ ID of AA of Sig Pep	First AA of AA Sig Secreted Portion	Last AA of AA Secreted Portion	Last AA of AA Secreted Portion
143	HOHCC74	209346 10/09/97	pCMV Sport 153 558 1 558	327 327	302 1	20	21	48
144	HPMFY57	209346 10/09/97	Uni-ZAP XR 154 1201 1 1201	250 250	303 1	30	31	42
145	HFXDN63	209346 10/09/97	Lambda ZAP II 155 1026 1 1026	33 33	304 1	14	15	53
146	HADCL76	209346 10/09/97	pSport 1 156 904 1 904	108 108	305 1	29	30	75
147	HMMAS76	209346 10/09/97	pSport 1 157 916 1 916	13 13	306 1	29	30	62
148	HMKCG09	209346 10/09/97	pSport 1 158 921 60 921	221 221	307 1	28	29	49

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

- 5 The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988).
10 Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

15 Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1
20 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra*.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

25 In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results
30 shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., +
35 or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

10 **Polynucleotide and Polypeptide Variants**

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

15 By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

20 As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization

Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the lenght of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query

amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions,

- 5 interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be

- 10 determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and
15 subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.
20

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity.

- 25 For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are
30 considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.
- 35

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired 5 residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another 10 example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C- termini of the subject sequence which are not matched/aligned with the query. In this 15 case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or 20 activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in 25 the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. 30 Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be 35 deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after

deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological
5 activity similar to that of the naturally occurring protein. For example, Gayle and
coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational
analysis of human cytokine IL-1 α . They used random mutagenesis to generate over
3,500 individual IL-1 α mutants that averaged 2.5 amino acid changes per variant over
the entire length of the molecule. Multiple mutations were examined at every possible
10 amino acid position. The investigators found that "[m]ost of the molecule could be
altered with little effect on either [binding or biological activity]." (See, Abstract.) In
fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide
sequences examined, produced a protein that significantly differed in activity from wild-
type.

15 Furthermore, even if deleting one or more amino acids from the N-terminus or
C-terminus of a polypeptide results in modification or loss of one or more biological
functions, other biological activities may still be retained. For example, the ability of a
deletion variant to induce and/or to bind antibodies which recognize the secreted form
will likely be retained when less than the majority of the residues of the secreted form
20 are removed from the N-terminus or C-terminus. Whether a particular polypeptide
lacking N- or C-terminal residues of a protein retains such immunogenic activities can
readily be determined by routine methods described herein and otherwise known in the
art.

Thus, the invention further includes polypeptide variants which show
25 substantial biological activity. Such variants include deletions, insertions, inversions,
repeats, and substitutions selected according to general rules known in the art so as
have little effect on activity. For example, guidance concerning how to make
phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al.,
Science 247:1306-1310 (1990), wherein the authors indicate that there are two main
30 strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural
selection during the process of evolution. By comparing amino acid sequences in
different species, conserved amino acids can be identified. These conserved amino
acids are likely important for protein function. In contrast, the amino acid positions
35 where substitutions have been tolerated by natural selection indicates that these
positions are not critical for protein function. Thus, positions tolerating amino acid
substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham 5 and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the 10 protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues 15 Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, 20 where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino 25 acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins 30 with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

35

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, 5 and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 10 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 15 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. 20 Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the 25 deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding 30 region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the 35 mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-

60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred.

- 5 Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-

- 10 forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide
15 fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

20

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

25

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., *supra*; Wilson et al., *supra*; Chow, M. et al., *Proc. Natl. Acad. Sci. USA* 82:910-914; and Bittle, F. J. et al., *J. Gen. Virol.* 66:2347-2354 (1985).) A preferred immunogenic epitope includes 5 the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a 10 denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from 15 the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., *J. Nucl. Med.* 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and 20 humanized antibodies.

20

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the 25 polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention 30 include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of 35 the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the

polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and 5 specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of 10 mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

15 Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, 20 deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. 25 Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, 30 such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope 35 derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *Basic Methods 5 In Molecular Biology* (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, 10 phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also 15 be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production 20 procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the 25 translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

30 Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome 35 identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat

polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be
5 selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the
10 polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome
15 specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence *in situ* hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al.,
20 "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides
25 correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage
30 analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease
35 could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined.

First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the
5 mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected
10 individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene
15 expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science
20 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model
25 systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the
30 present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of
35 restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying

personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

5 The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set
10 of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as
15 tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more
20 restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of
25 unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

30 In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using
35 DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected *in vivo* by imaging. Antibody labels or markers for *in vivo* imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. *In vivo* tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the

proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders 5 may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in 10 treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: 15 blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also 20 be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet 25 disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in 30 treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the 35 present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

- Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, 5 Myasthenia Gravis, Neuropathy, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.
- 10 Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.
- A polynucleotide or polypeptide of the present invention may also be used to 15 treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits 20 an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.
- Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory 25 response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel 30 disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative 35 disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect

interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, 5 or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

10 Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, 15 pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary 20 Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

25 A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the 30 polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following 35 DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes

- Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., 5 Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, 10 Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.
- Similarly, bacterial or fungal agents that can cause disease or symptoms and that 15 can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia). Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, 20 Coccidioidomycosis, Cryptococcosis. Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, 25 and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prostheses-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. 30 A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, 5 Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide 10 of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide 15 of the present invention, and returning the engineered cells to the patient (*ex vivo* therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

20 A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal 25 disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and 30 skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase 35 regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue

regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate 5 nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized 10 neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

15 **Chemotaxis**

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial 20 cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular 25 trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to 30 tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present 35 invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding 35 of the polypeptide and the molecule may activate (agonist), increase, inhibit

(antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying 5 agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity , and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

10 A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

15 A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

20 A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

25 A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

30 **Other Preferred Embodiments**

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

35 Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of

positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous 5 nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of 10 contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

15 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide 20 sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a 25 nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

30 Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

35 Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type

Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous 5 nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete 10 open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising 15 a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

20 A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone 25 identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% 30 identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of 35 comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in
5 Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as
10 defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at
15 least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method
20 comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino
25 acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an
30 amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is
35 performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide 5 comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

10 Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

15 Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid 20 sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human 25 cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an 30 individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of 35 illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

5 Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For
10 example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited Plasmid</u>
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
15	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSport 2.0	pCMVSport 2.0
	pCMVSport 3.0	pCMVSport 3.0
20	pCR®2.1	pCR®2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

35 Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain

DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into *E. coli* strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ^{32}P - γ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction 5 is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are 10 performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding 15 portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids 20 Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the 25 desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged 30 RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

35 This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is

used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

5

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X.,
10 according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by,
15 among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is
20 then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are
25 mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on
35

either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

5 **Example 5: Bacterial Expression of a Polypeptide**

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^R), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^R). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high

affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed 5 with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The 10 recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer 15 plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a 20 neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

25 DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or 30 Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

35 **Example 6: Purification of a Polypeptide from an Inclusion Body**

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem

columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

15 **Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System**

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcLM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring

signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," 5 Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

10 The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

15 The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

20 Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a 25 microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then 30 incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

35 After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture

and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the 10 recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ³⁵S-methionine and 5 µCi ³⁵S-cysteine (available from Amersham) are added. The cells are 15 further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

20 **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional 25 elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

30 Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used 35 include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

5 The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and 10 Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a 15 chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse 25 DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol 30 outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially 35 available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 µM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., *Nature* 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

5 For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with 10 BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that 15 the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a 15 heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACACACATGCCACC GTGCC
CAGCACCTGAATTGAGGGTGCACCGTCAGTCTCCTCTCCCCAAAACC

20 CAAGGACACCCTCATGATCTCCGGACTCCTGAGGTACATGCCGTGGTGGT
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGGAGGAGCAGTACAAC
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAAGGACTGGCTG
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC
25 ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCCACAGGT
GTACACCCTGCCCTCATCCGGATGAGCTGACCAAGAACCAAGGTACGCC
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATGCCGTGGAGTGGGA
GAGCAATGGCAGCCGGAGAACAAACTACAAGACCACGCCCTCCGTGCTGG
ACTCCGACGGCTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA
30 GGTGGCAGCAGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGTAAATGAGTGC
GACGGCCCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

35 The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera

containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

5 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., *Nature* 256:495 (1975); Köhler et al., *Eur. J. Immunol.* 6:511 (1976); Köhler et al., *Eur. J. Immunol.* 6:292 (1976); Hammerling et al., in: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

10 The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (*Gastroenterology* 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

15 Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, 5 secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies 10 described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulian et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 15 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be 20 tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) 25 and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine 35 (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in

Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl₂ (anhyd); 0.00130 mg/L CuSO₄·5H₂O; 0.050 mg/L of Fe(NO₃)₃·9H₂O; 0.417 mg/L of FeSO₄·7H₂O; 311.80 mg/L of KCl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄·H₂O; 71.02 mg/L of Na₂HPO₄; .4320 mg/L of ZnSO₄·7H₂O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H₂O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine;

0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of
10 the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from
15 each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further
20 provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation
25 of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site “GAS” elements or interferon-sensitive responsive element (“ISRE”), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called
30 Signal Transducers and Activators of Transcription, or “STATs.” There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at
35 higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN- α , IFN- γ , and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u>	<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	<u>STATs</u>	<u>GAS(elements) or ISRE</u>
5	<u>IFN family</u>							
	IFN-a/B	+	+	-	-	-	1,2,3	ISRE
	IFN-g			+	+	-	1	GAS (IRF1>Lys6>IFP)
	Il-10	+	?	?	?	-	1,3	
10	<u>gp130 family</u>							
	IL-6 (Pleiotrophic)	+	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
	Il-11(Pleiotrophic)	?	+	?	?	?	1,3	
	OnM(Pleiotrophic)	?	+	+	+	?	1,3	
	LIF(Pleiotrophic)	?	+	+	+	?	1,3	
15	CNTF(Pleiotrophic)	-/+	+	+	+	?	1,3	
	G-CSF(Pleiotrophic)	?	+	?	?	?	1,3	
	IL-12(Pleiotrophic)	+	-	+	+	+	1,3	
20	<u>g-C family</u>							
	IL-2 (lymphocytes)	-	+	-	+	+	1,3,5	GAS
	IL-4 (lymph/myeloid)	-	+	-	+	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	+	5	GAS
	IL-13 (lymphocyte)	-	+	?	?	?	6	GAS
	IL-15	?	+	?	?	+	5	GAS
25	<u>gp140 family</u>							
	IL-3 (myeloid)	-	-	+	-	-	5	GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	-	5	GAS
30	GM-CSF (myeloid)	-	-	+	-	-	5	GAS
35	<u>Growth hormone family</u>							
	GH	?	-	+	-	-	5	
	PRL	?	+/-	+	-	-	1,3,5	
	EPO	?	-	+	-	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
40	<u>Receptor Tyrosine Kinases</u>							
	EGF	?	+	+	-	-	1,3	GAS (IRF1)
	PDGF	?	+	+	-	-	1,3	
	CSF-1	?	+	+	-	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., *Immunity* 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

5':GCGCCTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCCG
10 AAATGATTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTGGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCCGAAATG
20 ATTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCC
CTAACTCCGCCATCCGCCCTAACTCCGCCAGTTCCGCCATTCTCCGC
CCCATGGCTGACTAATTTTTATTATGCAGAGGCCGAGGCCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTGGAGGCCTAGGCTT
TGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, 30 alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS:SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using Sall and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfecants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final 5 concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

10 On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

15 Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

20 After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

25 The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

30 As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the 5 Jak-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfet U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^6 U937 cells and wash with 10 PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 1 mM MgCl_2 , and 675 uM CaCl_2 . Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the 30 protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are 35 activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon

activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or 5 differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

10 The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCGG -3' (SEQ ID NO:6)
5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

15 Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes Xhol/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

20 To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) 25 containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

30 Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

35 To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

(Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count
5 the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR
10 can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-
20 κ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target
25 genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in treating
30 diseases. For example, inhibitors of NF- κ B could be used to treat those diseases related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

5' 5':GCGGCCTCGAGGGACTTCCCGGGACTTCCGGGACTTCCGGGAC 3'
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

10 PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene) Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

15 5':CTCGAGGGACTTCCCGGGACTTCCGGGACTTCCGGGACTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCA
TCCCGCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACT
AATTTTTTTATTATGCAGAGGCCGAGGCCGCTCGGCCTTGAGCTATT
20 CAGAAGTAGTGAGGAGGCTTTGGAGGCCTAGGCTTTGCAAAAGCTT:
3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and HindIII.

25 However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the 30 NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described

in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

5 As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below:

10 Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 µl of 2.5x dilution buffer into Optiplates containing 35 µl of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

15 Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

20 Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25

28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Bioteck washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is

incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca⁺⁺ concentration.

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating

tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 5 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are 15 used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇ and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 25 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and 30 centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a 35 biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM

- 5 ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

- 10 The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide.

- 15 Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

- 20 Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

25 **Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity**

- As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1
5 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and
10 cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a
15 positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with
20 Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in
30 SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTHERM Polymerase. (Epicentre Technologies).
35 The intron-exon borders of selected exons is also determined and genomic PCR

products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 5 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigeninideoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and 15 propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and 20 chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

25

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is 30 a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with 35 specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10.

The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

5 The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

10 Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

15 Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

20 The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

25 As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If 30 given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending 35 on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

- The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules.
- Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as 5 ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, 10 manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

15 The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

20 Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

25 Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

30 The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

5 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the
10 presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

15 The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

20 Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media
25 from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

30 The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

35 It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other

disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

305

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>212</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>		
Name of depositary institution American Type Culture Collection ("ATCC")		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit <u>25 SEPTEMBER 1997</u>	Accession Number <u>209299</u>	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <u>Sonya Barnes</u> PCT International Division		Authorized officer <u>SBS</u>

306

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>215</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>		
Name of depositary institution American Type Culture Collection ("ATCC")		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit <u>25 SEPTEMBER 1997</u>	Accession Number <u>209300</u>	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		
<input checked="" type="checkbox"/> This sheet was received with the international application		
For International Bureau use only		
<input type="checkbox"/> This sheet was received by the International Bureau on:		
Authorized officer <u>Sonya Barnes</u> PCT International Division	Authorized officer	

307

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>220</u>, line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>
Name of depositary institution American Type Culture Collection ("ATCC")		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit 02 OCTOBER 1997	Accession Number	209324
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p>		
<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p>		
<p>Authorized officer <i>SDG</i> Sonya Barnes PCT International Division</p>		
<p>Authorized officer</p>		

308

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description
 on page 225, line N/A

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection ("ATCC")

Address of depositary institution (*including postal code and country*)

10801 University Boulevard
 Manassas, Virginia 20110-2209
 United States of America

Date of deposit

09 OCTOBER 1997

Accession Number

209346

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)This information is continued on an additional sheet D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

Sonya Barnes
 PCT International Division

Authorized officer

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

- (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
- (f) a polynucleotide which is a variant of SEQ ID NO:X;
- (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
- (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
- (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.

3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

- (g) a variant of SEQ ID NO:Y;
- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:
(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and
(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.

18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

- (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

- (a) contacting the polypeptide of claim 11 with a binding partner; and
- (b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

22. A method of identifying an activity in a biological assay, wherein the method comprises:

- (a) expressing SEQ ID NO:X in a cell;
- (b) isolating the supernatant;
- (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

<110> Human Genome Sciences, Inc.

<120> 148 Human Secreted Proteins

<130> PZ019, PCT

<150> 60/063,099

<151> 1997-10-24

<150> 60/063,086

<151> 1997-10-24

<150> 60/063,387

<151> 1997-10-24

<150> 60/063,148

<151> 1997-10-24

<150> 60/063,386

<151> 1997-10-24

<150> 60/062,784

<151> 1997-10-24

<150> 60/063,091

<151> 1997-10-24

<150> 60/063,090

<151> 1997-10-24

<150> 60/063,099

<151> 1997-10-24

<150> 60/063,092

<151> 1997-10-24

<150> 60/063,111

<151> 1997-10-24

<150> 60/063,101

<151> 1997-10-24

<150> 60/063,109

<151> 1997-10-24

<150> 60/063,110

<151> 1997-10-24

<150> 60/063,097

<151> 1997-10-24

<150> 60/063,100

<151> 1997-10-24

<150> 60/063,098

<151> 1997-10-24

<160> 492

<170> PatentIn Ver. 2.0

<210> 1

<211> 733

<212> DNA

<213> Homo sapiens

<400> 1

gggatccgga	gcccaaatct	tctgacaaaaa	ctcacacatg	cccaccgtgc	ccagcacctg	60
aattcgaggg	tgccaccgtca	gtttccctct	tccccccaaa	acccaaggac	accctcatga	120
tctccggac	tcctgaggtc	acatgcgtgg	tggggacgt	aagccacgaa	gaccctgagg	180
tcaagttcaa	ctggtaacgtg	gacggcgtgg	agggtcataa	tgccaaagaca	aagccgcggg	240
aggagcagta	caacagcacg	taccgtgtgg	tcagcgtcct	caccgtcctg	caccaggact	300
ggctgaatgg	caaggagtagc	aagtgcagg	tctccaacaa	agccctccca	accccccacatcg	360
agaaaaccat	ctccaaagcc	aaaggccagc	cccgagaacc	acaggtgtac	accctgcccc	420
catcccgaaa	tgagctgacc	aagaaccagg	tcagcctgac	ctgcctggtc	aaaggcttct	480
atccaagcga	catcgccgtg	gagtgggaga	gcaatgggca	gccggagaac	aactacaaga	540
ccacgcctcc	cgtgctggac	tccgacggct	csttcttcct	ctacaqcaag	ctcaccgtgg	600
acaagagcag	gtggcagcag	gggaacgtct	tsttcatgctc	cgtgatgcat	gaggctctgc	660
acaaccacta	cacgcagaag	agccctctccc	tgtctccggg	taaatgagtg	cgacggccgc	720
gactctagag	gat					733

<210> 2

<211> 5

<212> PRT

<213> Homo sapiens

<220>

<221> Site

<222> (3)

<223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2

Trp Ser Xaa Trp Ser

1

5

<210> 3

<211> 86

<212> DNA

<213> Homo sapiens

<400> 3

gcgcctcgag	atttcccgaa	aatctagatt	tcccccgaat	gatttcccg	aaatgatttc	60
cccgaaatat	ctgcctatctc	aattag				86

<210> 4

<211> 27

<212> DNA

<213> Homo sapiens

<400> 4

gcggcaagct ttttgc当地 cctaggc

27

<210> 5

<211> 271

<212> DNA
<213> Homo sapiens

<400> 5
ctcgagat~~t~~ cccccaaatc tagatttccc cgaaaatgatt tccccc~~aa~~at gatttccccg 60
aaatatctgc catctcaatt agtcagcaac catagtccc~~cc~~cctaactc cgccccatccc 120
gccccctaact ccgccc~~tt~~ttt~~gg~~actt cccat ggctgactaa ttttttttat 180
ttatgcagag gccgaggccg cctcggccts tgagctattc cagaagtagt gaggaggctt 240
ttttggag~~gc~~ ctaggcttt gcaaaaaagct t 271

<210> 6
<211> 32
<212> DNA
<213> Homo sapiens

<400> 6
g~~cg~~ctcgagg aatgacagcg atagaacccc gg 32

<210> 7
<211> 31
<212> DNA
<213> Homo sapiens

<400> 7
gcgaagcttc g~~cg~~actcccc ggatccgct c 31

<210> 8
<211> 12
<212> DNA
<213> Homo sapiens

<400> 8
ggggactttc cc 12

<210> 9
<211> 73
<212> DNA
<213> Homo sapiens

<400> 9
gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60
ccatctcaat tag 73

<210> 10
<211> 256
<212> DNA
<213> Homo sapiens

<400> 10
ctcgagg~~gg~~ga ct~~tt~~ccccgg gactttccg~~g~~ ggactttcc~~g~~ ggactttcca tctgccatct 60
caattagtca gcaaccataq tcccgcccc~~t~~ aactccgccc atcccgcccc taactccgccc 120
cagttccgccc cattctccgc cccatggctg actaattttt tttattat~~g~~ cagaggccga 180
ggccgcctcg gcctctgagc tattccagaa gtat~~gt~~gagga ggctttttg gaggcctagg 240

cttttgcaaa aagctt

256

<210> 11
<211> 826
<212> DNA
<213> Homo sapiens

<400> 11

ggcacgagaa tataatggct	tttgtgattc tatttatagg	aatgtccag aataggcaaa	60
tctacagaaa cagaaaaacta	gattggtgat cgccctagagc	ttggggcagg ggggtgggag	120
tggtaatga cttctaattga	gtttttttt aaggtatga	aatatgttga aaattgattg	180
tgattattgt actaaaaacc	atthaacgta tattaagggt	ggttaattgt atggcatgcg	240
atttatatct caacaaagct	gtgagtgtgt ggccttcatgt	atggatgtgt atgtgtgtgt	300
atataatctt atacatgtat	acatggatgc ccatgtgtat	ctatgttagaa tatgtaaaac	360
aaacatgagg tagtttgata	ttttagtctg gagctacaga	gagatataag cccagcgatc	420
aagattcaga aatcagcagt	cactgagggt gtatgtatgt	atgggattgt ctaaaggaaa	480
tqagagggga ggagaatggg	tttccacaga caaccctttt	tggaaactga aaagaaatca	540
ctagacagaa gggaaatgaac	tagagaagac tggcttaactt	ggaggtcaag tgtgagcttc	600
attttctgcc tgcgagggtgg	gaaacttatt tttaactgtat	tctctgggtt tcaacacatc	660
ctctgggttc tccagggcat	aggggaggtgc tgctgtgtca	ctgtgcactgt ggggagtgga	720
agactgaata aatattgcaa	atggagggac cagccagagg	gtgcaaaggc ctggggagca	780
tgagggaatg cagctcacca	gcagagtttc aagcagtta	ctatgt	826

<210> 12
<211> 524
<212> DNA
<213> Homo sapiens

<400> 12

gcacagaggg cttgggtgca	ggtggtttat tgggaagtc	atccctggaaa atccaaaagg	60
aagggatgga gaagagatag	aagacaagaa agaatgcatt	gctcggttgt catgggtata	120
aaaagtttctt aggaagcttc	tgcagaaccc tatgcaatgt	gcctcgaaatt gtccaaaggaa	180
ttgaatgggg agctggtgca	tttgcacact asttctgttg	ctcaactgtg ggcaacagg	240
cttttatccc cagctttcc	aggctgeccc ggggagacag	cagctatggg gaggcaccaa	300
cccatgggt gtactcattc	cagaatcctt cttcccttac	acgctgacag tcaattattc	360
accaagttgt aacttcgaat	tctacttacc taaaatgcgt	ttggcataca tctgcatgtc	420
acactcacac tgtccctatc	ttggtcgaga cattataatc	actctctgtg actactgcag	480
cagttcccta gctgaactcc	tggctcatct ggtctatatt	gctg	524

<210> 13
<211> 491
<212> DNA
<213> Homo sapiens

<400> 13

ggcacgaggc aaaagcttgt	gctgttagct taaaagtgtat	tttaaaaataa atctgaaatc	60
atttaaacag catgaacctt	ggtggccaaa tagatcaatg	acaaagagga gaaaacctag	120
atacaggttc attttgcct	tatatgtttt gagatttagt	tttctattt ggtctgtgac	180
taatacagat gcatcacggc	tgagagcaaa gcgggtgaa	tgtccctatt aattgccacc	240
atgggtgcag gctggaaatga	gggtgtggcc agctaagagg	ggatttgctc ttcttgccct	300
agaagttctt cattgtttcc	tgtctgtct tggccac	tgttagcac acttcctttt	360
gttattttat gcttttata	gctggaaaccc tgagggttcc	cagaatctg cacatgctt	420
ctagatggtg ctctggattt	tcttaaaga taggaagaaa	aaggcaaagg caggtctgt	480
acgcttctta c			491

<210> 14
<211> 403
<212> DNA
<213> Homo sapiens

<400> 14
gqcacgagcg gggccctaga gagcaactcg³ aggtccaaac ccctcatcct aaagaagggg 60
acgcgtgcggc catgacattt catgcctccc aaggctctag agctataaaa tggtgagcc 120
atgactgagg gcctgctgtc ttctcttc³ ttgttactgt atttattaac ctggttactt 180
atgcttcca aaaagcttta tgtgcaa³ atcttttgct ataatccaca cttcagtcag 240
atggatgcat gcaatggaac cagtcagaag atccacaatg ctagacagt³ cacctgatgt 300
gcagttcctg gaatggagct ctctcc³ aagccaaatg ttttctctga aaccctctgt 360
tctttaacgc tgaagtcttg gatgcctgct aggagcagct cga 403

<210> 15
<211> 813
<212> DNA
<213> Homo sapiens

<400> 15
ggcacgadat cattttctgt cccctccat cttaggctga ccggttccct gatgtgttac 60
ctgtttctgc tactgatcca aactgcagaa ctttcattt atccccagg cctccaggca 120
gtatccaatg gggaatcgc tctaaaagg³ accagaccaa cggttccag ccccttattt 180
ctggtgactg aggggaggaa agaatgggag ggggtattct tgcgttgcgttgg atggaaagga 240
aacacactgt caaattacta tatctccctt³ gtttttattt acagtagaaat tctccagcca 300
tatttttattt gtctatgggg gaagttggag atgggtgacct tgattagaag tgcgttggagg 360
gggataaaatg gaggggataa gatttcagtt ggttttgaa aatgttaaag tcttaaaaata 420
atgcgtccca tctgaagaat tttttctaaa accagagttt ataaaaatata cactgataca 480
gcctgcccccc tcattttccctt gccacaggag atgttttgg³ ctagagacat ttgtttaata 540
ataqcttgc tctgatattt ccagtagctt ccctctgtgt³ gaggaaagga tagaaatgtt 600
caggacatca tcatacaggc tcctcatctt³ caaagttcca gtgcgttga cgcctacacg 660
gaagacttgg aactgcaaaac aggctgggg³ cacctcagtg acatctgacg ctgtccaaacc 720
agaagttcga tttttgttct ggggggtgaag gaggaaacag actgtactaa aggactaaaa 780
taatttgc tatactaaaaa aaaaaaaaaaaa aaaa 813

<210> 16
<211> 264
<212> DNA
<213> Homo sapiens

<400> 16
gtttaatata tgcgttgcgttgg attattaagt atgacattct cttttcttct agagttttgt 60
tcagttggcc³ aaaggctaaatg attagcagat gttttagaaatctatgcagga tattttaaaa 120
tggtttagtg actataccctt gagggcagat ataaatgtt caagagattt ataggaaag 180
gatttcttgc aagatttgcgttgc agatggggcc³ aaatgttgcgttgc agctctcaat agctatccaa 240
aacacctgg³ atgtttcttc tcga 264

<210> 17
<211> 520
<212> DNA
<213> Homo sapiens

<400> 17
ggagaaggact ttatgcaggg aagtgcacgc³ ggacacggag ggactcataat ttaccgagct 60

ttggtgtcagt ggccccggc ctgggtattc tatttaagcc atgcaaaaac ccattggga	120
gaagagtaa ggttttcctt ccgcaggaaa aacttgaggc tcagagaggc tatgagacat	180
gagacatgcc aggtcacaca gctggtagct ggcaaaqctg actccaacct gtgtctgagg	240
gactctgaaa cctggttctg gcccccaactc tgggcagccct gctccctctt acaagccact	300
gcctgcagat taagcagtc tagcaaaggc ctgggagcat ccagagagtg cccctggctg	360
gcgagtggtt gagcagccct ggtttccctt ctggaccct caaggatcac aggagtgtca	420
cccagaagta acttaactta tgagtgtttt atgaacagga aaagcaggaa aaggggtaaa	480
gtcacatgtt ttccacaacca aacagcctgt aaactcgtgc	520

<210> 18
<211> 993
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (474)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (478)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (551)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (581)
<223> n equals a,t,g, or c

<400> 18
ggcacgagga acaaactagg gtagaaggcc tggggccggc agactgatca tgtgatgctt 60
agccttatcta ctattctctc taatcctgct agtagtagat ctgccagtgt tccatcttag 120
ctgatgttac tgggtctagg ggagtgaccc ttgtggctt tctagccatt agccctgtta 180
ggatagactg gtttgtactg gtgcagttctg tttaaaaaca tgcatttctc atggctcata 240
atgtttgtgt gttaagtctc gccatgcagt ggtcttaatc ttgttttcaa ctggtagctc 300
tgccataacct gtctctctaa ggagacctaa ctattgccta ctccccaccc gcggccattc 360
atccaccagg cccaaacctt tgggtgttga acaatacagc tacttatttt tgacatgtgt 420
ctttatgtgt gtgtcacttc agtggaaagtc aacccaaccc tggtaggag acantacntg 480
tatgaggaaa gggatcacag gtacagaagt tcacagaact aatgcacttt tcacatttt 540
gtgctcataa ngcatttccs cctatagata tgatttgaga nagaagacac tgaaagaatg 600
gaggaataga caccaagtta atarggttc ctaactgatg aatttcactc ttaggatggc 660
tgagccagag accaccattt cattctttt gctgtgcctt gcctgtttcg atggtttcc 720
aggattccca acgtgataag tgggtccctt gttgacgtt tttatctat tctggcaatt 780
cagtgtcagt atccgttttc ctttcaataa ttccaccaaa gtatttatgc cccactactg 840
tatttcattt attgttaatg acacgctggt ttttatttgc atttcgttgaag ttgaaatgctg 900
tggAACATTG GTGGTGTGTC ATAATACCTT AGATTGTT ACTTAGTGTT ACACCAAATA 960
atggtaggtt gatgggtttt tagattgcct cga 993

<210> 19
<211> 459
<212> DNA

<213> Homo sapiens

<400> 19

ggcacgagga	agcgtgaacc	ccagggaaaca	gggggtccct	tccctcctca	gacacaagcc	60
acctcagctt	gtggcttttg	ccccccagcc	ccaccaaccc	acctgttcat	ttattcaaca	120
gacaatgaca	gctgatattt	attggacatt	tgcaccaatgc	caaqcattcg	gttggat	180
tcccatttgt	ttctcacagc	cgttatttat	tgtctgtcc	tctgtgccag	gtgtgtgt	240
ctggggcaggg	gcactgcatg	ggctgcctgc	cstggtggag	cttgggtct	gatgggtgag	300
gctgacccaa	gcccacccca	ttgccaacag	gcccaggcga	agagtacaca	caggggcctc	360
ataccatatg	tctaaatatt	taaaaagtta	tsaatcaagc	taacaactgt	taaataaaaat	420
atgttctatt	ctcctacttt	aaaaaaaaaa	aaaaaaaaaa			459

<210> 20

<211> 555

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (50)

<223> n equals a,t,g, or c

<400> 20

agctccaccc	cgttggcgcc	cgctctagaa	cttagtggat	ccccccgnch	tgccaggaat	60
tccgccacga	gcaaggtcac	agagctagaa	aaaggcagaa	ttgggaccta	tacccagaat	120
ttcttaactct	agctctgtaa	agctggaca	ctggagaagc	agaggttttg	gtgttagtact	180
cctctgagct	cgggtctta	gaagtccaca	tggggctgt	ggagtggta	ggggagatgg	240
aggtgggaag	aagggagaag	acccctattc	ccctattctc	tttcaatcag	agaggattcc	300
tcgacttatt	tacctgcctg	tgtatctate	tgaagagaat	tctatgaccc	cccaaaatct	360
gcgatttcact	ctgttccagt	tctgttactc	tttatatctg	gagctagaac	ttggctttag	420
atcaactgtca	caagaggtga	ccagagaatg	ggcttgagt	tacttcttt	taataaaaat	480
ttgctggcaa	gttcctgtga	gtgagttccc	gtttgtaaaa	gagaacccat	tcttacttct	540
ggagaaaaaa	ctcga					555

<210> 21

<211> 665

<212> DNA

<213> Homo sapiens

<400> 21

ggcacgagaa	actccagtt	atgccattt	ttttgcctt	accccccgtc	60
cttctagggg	ttataatgag	aagaaactaa	cagacaata	tcaatgttag	120
ttctttatga	agaagccaca	gcattcctaca	aggaagaaat	cgtgcattcag	180
ataaacccaga	agagctagaa	aataatgtag	atcagatctt	gaaatggatt	240
tcaaagatca	taactcttga	cttataaggc	tagctactt	ataatcactc	300
ctctgccgac	atcatagaaa	ttgttcaagt	gtcagtaaca	ctttattaaa	360
agaaccagca	ggtggatagt	atataaggttt	atgcctgtgt	atctttctc	420
ctaaacatga	aatataatga	atatagttaat	tattaaggga	catgagaaag	480
ttaatactta	aattgctaaa	gaataaaataa	atctgacaaa	atgggtggat	540
tttattacag	aaaaaaaaatgc	agatgatctc	ttaaaaataaa	actaaagata	600
aaaaaaaaaa	aaaactcgta	gggggggcyc	cggtacccaa	tcgcctatg	560

attac

665

<210> 22
<211> 777
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (274)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (278)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (295)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (676)
<223> n equals a,t,g, or c

<400> 22

ggcagagctc	aggtaagarg	caaaaattact	agaatattca	ctctcactga	aaatgagtaa	60
aaacctaact	tagataaaaa	tccttatctt	gttcattttt	atccctggcc	ttttgggttg	120
gaagaatggg	ccagaccatg	tgtgtgtgt	tatgtgtgt	cgtgtgtgt	tgtgtgcgca	180
cttgggttta	tttatatgg	ccggtaaaat	ttcgttcacc	attaattttat	gttaatttac	240
caacctctta	aatgagaaca	gtgagaattt	tctncatngt	taataataca	ctggncagtg	300
catatatgca	tcacgaagag	aggattttcc	cattgataat	agatttccaa	atacatcttc	360
ctgcttaag	attttaatat	atggatttat	atataaaaac	tagtaagtgc	attggaaaag	420
caaactctca	wccttctctt	atttgagawc	tcaacttttaq	aaagtctatg	ttctcaacta	480
cagaaaaataa	tttttagacc	agctaacttt	cagatttctg	cagtgtttat	tttctcccag	540
ttgagggttg	gtttttgtt	gtttgttgt	ttgtttgtt	ttctgtatta	aaaagtaaga	600
atacggccag	gcbcgcatagc	tcatgcctt	aatcccagea	ttttgggagg	ccgaggagg	660
cagatcacct	gaggtncagg	agttcgagac	cagcctggct	aacatggtga	aacccagttt	720
ctactaaaaa	aaaaaaaaaa	aaacttcgag	ggggggtccc	ggtaccta	cgtccct	777

<210> 23
<211> 540
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (341)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (378)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (425)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (450)
<223> n equals a,t,g, or c

<400> 23

gaggatcccc acagggcccc tctgttagcc	tggggagtcg gcagtgc	tctaggcccc	60
tttaggagagg gggcaggggg	gcagcagtag aaatgtggcg	gggtcggact	120
ggccgtcttt gtgtttrtgt	tgttatgtg gagtgtcatt	gggtctttat	180
gtttcaagtct ctccatgtgt	caggtctgc ctggctgtcc	cttggtgtatt	240
ccatctgtct agccctgtgt	tccatgtca	ttctctggam	300
catctggggc atctgttttg	ttgggtttgt	tttgggtgtca	360
tctctgtcc ccgagcanct	gtctctgtcc	tgtgagtc	420
tcganttggc cccaaaccaag	ccccc	tctgcctgcc	480
tctttccaa cccctccaaa	tgaqatgtt	gagtgcgtg	540

<210> 24

<211> 484
<212> DNA
<213> Homo sapiens

<400> 24

ggcacgagg	ccggggaccg	agggatgtga	gcctcggtta	caactccagg	acaggaggga	60
gagaatgcaa	ctcagcctgt	ccctctgtgc	atttgtggta	tgcactaacg	ctgtctgcac	120
acatgcagct	accaaccaag	ccagactgg	ggggttcata	aaggctctga	ggcccggcca	180
cagccccctt	tgcctttagg	ttgcttccc	ccccagttgc	ccgcctctgg	tatttgttgc	240
tacttacaga	atcttttagga	ccaaagggt	gagcgtgggg	ccaaagaatct	ggtgagcaac	300
aagtcaactgg	ctctggccct	ccacttttg	acaggggtgt	cccggggtgg	ggagacggtg	360
atccttagacc	cgctgtcacc	tgtggggctg	ttcaagtgc	tgagggtaaa	gaacgagttg	420
gtcccactgc	tcatagttt	ccctgcac	ttggcacagg	gcatagcaca	aagcaagccc	480
tcga						484

<210> 25

<211> 707
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (562)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (570)
<223> n equals a,t,g, or c

<400> 25

gggtcgaccc	acgcgtccgc	ccacgcgtcc	gtttctaca	acccttagga	acatcagaat	60
catgtgtgtg	tgggtgttta	ttaaataaaam	sagttcctgg	agtcactcc	cagtgactgc	120

cagtctgatg attagggct cagctaggac cttaggttgc gaaagctccc agctgatctc	180
atgcagccag cctggctctg gctctggckc tgggagctgg gttgggaact agtctttgg	240
gctattctgc tgawacttca agatggctc ttgactccg tcttgatttgc tcakcacttg	300
tattcaggctc tgttcttccc ctggatttgc aactccctga tgcctgggtc atctcagctc	360
atgagctgag cttwcagtgg gtgcctcagtg gaacagatgc tgaatggagt caggctgttag	420
ggaggccagc gtgtgttggg aagtggagaga caaaaatcat tttaaaaaga atcttttgc	480
ccttcagttt tgtttgcatt gagttaatgt gatctactct agtggaaagcc agtgcagctt	540
aagtggaggt cttgccttga antggagccn ggttatggat cagcagagct gccaaaagcg	600
ttttggggga aatgtttctg tgcaccctc agttgattga actcaagttt tcactccgt	660
ttaacaccac gtgggggcca ttctgacttc tgccggagtgg gtatgtat	707

<210> 26

<211> 793

<212> DNA

<213> Homo sapiens

<400> 26

ggcacgagta ccatttatttgc agctctagac acataactagg cagtgtgcta agacttgatg	60
tgtgcatgag ctcaccaaataa agtctgttgc tactatttgc tgtaatctgtt agggaaagcta	120
cataacttcc tcaagctatg twtttcttaac tcaaagacag gtaccttaaa gatcatctgc	180
tgctgcctct ctttgaccct actaacacag agtacaggtt attaagcaat ataaaacaag	240
ggttatttaac agactctgtt ttwgctgca tccccytcaa cacccccgag ctccacaaat	300
cttaaagtca aatgctcttgc gcttacttta ttgggttaca cacccctcagaa acgaacagaa	360
ctgtcaaaaca cctgtgaagg caaaagatca gcttgccttgc ctctacacagg cctctgcagc	420
agcttagcagt acttagctct cagtagcacc cactccaaaca gcaagggttt tagctgttcc	480
ttatacacac gcacacacat acgcacacatc acacacacac agactctgaa gcttccctgg	540
cctctccatc taactacca catccttacc tcttgcgttca aaaacttca actgggcttc	600
ccacttcaat actttctcaa ctcaaaaagg caagtgttctt tttaaacct taaatcagaa	660
cacgccactc ttctgcttca gattccccaa ggatttctac gcacttcata tctaaactac	720
ttacgatgac ccaaggccct actagatttgc gcttgccttac ttccagcgg cacgagagag	780
aactagtctc gta	793

<210> 27

<211> 638

<212> DNA

<213> Homo sapiens

<400> 27

gataagaaat tatttaaattt tctttatgaa tattgttccct caatttagcg ttcttcctca	60
ttttgcctat ttctctttta ttatccata tgggctgtt ttgttccagec aaacaatttgc	120
tagcatgtct gtttcaaaag taaaatagtgc atatattttaa agttctaaat gtgttcttta	180
tgtatttta aaggagatgg gtaaaaataga atgtatccctt ctttaccctg atgacattcc	240
cgtgatatat ttcaaataat attttgattt gggtaagcca gtaggaccaa atccatggtg	300
atcacagata cagattcaca aatgcata gagaatcata aatagatgca tatggaggag	360
tctgacagta tagtggaaattt ggtttcaagt aatttgcacat attagaacctt tcaggcatc	420
acctggccagt aatccttattt agaaatagga ttggaaatattt ggggtcacca gctcaagacc	480
attttttgtt gagagctgaa caataaccaa aagtcagagc tataggaata aaaatgaacc	540
tattccagtc attagaacttgc ttctctgttca taagctctttt ctccctctcc ttcataaaaaa	600
aaaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaa aactcgaa	638

<210> 28

<211> 528

<212> DNA

<213> Homo sapiens

```
<220>
<221> SITE
<222> (421)
<223> n equals a,t,g, or c
```

```
<220>
<221> SITE
<222> (436)
<223> n equals a,t,g, or c
```

```
<220>
<221> SITE
<222> (459)
<223> n equals a,t,g, or c
```

<210> 29
<211> 919
<212> DNA
<213> *Homo sapiens*

```
<220>
<221> SITE
<222> (380)
<223> n equals a,t,g, or c
```

```
<220>
<221> SITE
<222> (432)
<223> n equals a,t,g, or c
```

```

<400> 29
gatcgttttt gccattgcag tgaccaaacag gactgtggac ctcaagtaaag gctttcccta 60
catcagcatac tgcrgatcct tccccccctca ragctgcata ttcaagccagg tgctcaaatat 120
gggagctgct ctggccgcgt ggatctgcata tgtccgttac caccagctcc gggactgggg 180
cgtcagaagg tggcctaacc agctgatcct atggacgggt ctctctgtgtg ccctgggcac 240
ytccgtggta ggcattttac caggtgagac ccagtcggcg cccagggtct gtwmccggcc 300
ggcstytgga aytacaactc ccagcatgcc ccgtggccat aggcttwawg tctcgggggc 360
tggttcccgc cccgccttcen tgggacttgt atttttctct ggcatttgcc ctggaccggc 420
tggatccctt gntctctgag tggggcatta cgagcggagc tggggatgtatg taggattcgt 480
tgattccagg acgttgggat aattttctgc cagccccctct ccccagetta tttaatgtatg 540
aaattactgg tccaggcgcga gtggctcatg cctgttaattc cagcacatttgc ggaggcggag 600
gcaggcggat cgccctgaggt gaggagtttgc agaccagcct ggcaccaaccatggtaaaa 660
ccccgtctct cctaaaaatat gcaaaaatttgc gcccggcatg gtggcaggcg acttaatccc 720
agctacgtgg gaggcagagg cgggagaatc atttgaacct gggaggtggat ggttgcagtgc 780
agccgagatc gagccattgc actcaaacct gggggataag agtggagactt ctctcaaaaaa 840
aaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaa actcgttaggg ggggkycggg tacccaaacgc 900

```

gccctatagt ggatgcgtg

919

<210> 30
<211> 864
<212> DNA
<213> Homo sapiens

<400> 30

gctcggtgccg	caacatacta	catttatatt	aaagtgtatt	tggataataa	ttgccttct	60
tgtatgttg	attgttgcct	ttaagattti	aatatagaca	attaatagaa	ttgcttatgc	120
acatttcat	aatgggttg	ttgggttttg	tttattttgt	tataccttc	cttgcacatt	180
tgtgtccaaa	attacattt	ctcatataaa	atcatttgc	tgcagtttt	tcattatata	240
gtccgtaaaa	tacagatatt	tgtctctta	tgaatataaa	cattctctta	tcatttagatt	300
agattaaatg	agtgtatctt	ckgtaattt	taagttaawt	ggatttaaca	ttttgttgac	360
aaaacaccta	ggcaccagca	gttttggta	gccccatagt	ttagttgaa	gtagctagaa	420
tcctcttagt	tacagtttga	cgagtttcat	ctcacctatt	taggttttt	gttgggttga	480
tacttgacat	caaaggaaa	gcacccccc	tcttgagtga	cttcaaggat	gcattaaggcc	540
tgcagtgcct	ggcctcgatt	ttttccat	actgtgcctg	tatgtttct	gtaatcactt	600
ttggaggggct	gcttggagaa	gctacagaag	gcagaatagt	gagtacaaag	attggtagtg	660
gccaggctt	tagcttca	gaggcaagtg	tctgtatgca	tttgtctcac	tattcatact	720
tttattttgaa	gagtctaccc	acaacatgat	taacgtgacc	caaagcagac	tttccccaaa	780
gttaattgct	gtggaaaaca	tgggaagcc	atttgaacag	aagatgcaca	gttgaggtaa	840
aaaaaaaaaaa	aaaaaaaaactc	gwag				864

<210> 31
<211> 919
<212> DNA
<213> Homo sapiens

<400> 31

ggcacgagt	atgctgctgg	ctgccttcct	cgcctgttt	cctctccatg	actccagggg	60
tttgaagcac	acaggagctg	gacatgtcaa	tctgttagt	cctctccaa	taccactgaa	120
ggccgtgagc	ctctctcctg	tttccagcct	gcaggtgecc	tgtgtgtgt	cttcattcca	180
gtttctccctc	actttctct	cagtctcttg	agcttggaaq	ccttaactgta	gtttgtgtct	240
cctccctggg	cacttgaggt	caggctttt	ccttttgc	acatttgagcc	acatgcctt	300
gatacacagt	tgttagcaaag	aagggaggtg	atgaacttgg	ctcactttt	tttctgatc	360
ccctccctac	tcatccgtca	ctccccaccg	aaccccagat	atcttatagt	cctaaggctt	420
gttagaggatt	aaggaaagga	attggagatg	ggtttactt	agttcacaga	aaagctttct	480
ttgggatttt	tcctccccct	tagggcttt	taagtctagg	tgaagtgaaa	gttcacacat	540
gtgtttgtt	ggttgctctg	taattagcta	ctagttttt	tccctagacc	ttctctgctc	600
cagtgtcttg	ttcatgtgtc	ctgaccctgt	gtccttgaat	tcccacttt	ctttgggatt	660
taagttattt	tatgttgc	acaatattt	aagatgaaaa	agtcctgaag	aaaacttacc	720
aggttcttc	ctttggcttt	ttttttttt	tcttgcagg	tactgtaaat	tgttaacttag	780
ggatgccaag	caggcttgg	tcaatggcta	aacctttat	tgtattacag	tgtaatgtcg	840
atctcagcct	ggtctcaatg	ccagagcaca	cagagactt	aataaaactg	ttataacgat	900
aaaaaaaaaaa	aaaaaaaaaaa					919

<210> 32
<211> 956
<212> DNA
<213> Homo sapiens

<400> 32

ggcattttca	gtcaaaaata	tggctgccgt	ttaaggtgtg	agcttttgc	ctttttacct	60
agaaaaacaa	tgatatgtaa	atttcttatt	ataatttgc	ttactctact	cttatttgc	120

atttgtcaac tctgcaagag acaagggttg gtacagaaaa tatcattta tcagaaggaa	180
accttgcctt ctacagttag tactaccctt ttgagctgat tctggataat aaagcttgc	240
tcgcaataat ccagagttt ccaaaccata tatttgcctt cttttactga ctgatatgca	300
aagttgtgtt taactatatg gaaaattgtg aatcccttcc tcccaagtct acacactcca	360
cttgcttttc ctttctaccc tgaattcatg tgcatcccc cagtttctg cctttgtaat	420
ggaggcctca gttcttctgc agccacagtt gcaggaaaco caattgtaat cagcagctgc	480
cctgctgsta aaagctatta gtgccmatgt tggtaatgac cgaccaggat gaatgtgctg	540
attnaagggg tagacttatt tcttgcact ccctaggcc ttgttcttaa atgagattt	600
tacgttgttt agttgtttac tcctctagca taaggaggat taaccactaa ccaccacatt	660
cgatatacgac cacgttgaat taaaataatg gggaggtaa gacataatcac tacctttgag	720
gtatgaagac acgcgcctg aaaaaacttg gcttcaccc caatttttg cctaaagaag	780
ttggataccca ctgcaatttg ctgtgaagtt atacatgtta tttgttccaa gggactata	840
gatgagaagg gtacagttag gtttctttt agaactaaac acattagcag taacctgcaa	900
aaatcaatac caattacttg caagcaaggg cttaaaaaaaaaa actcga	956

<210> 33

<211> 566

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (400)

<223> n equals a,t,g, or c

<400> 33

gggttcctga gctgactagg taggttagtga gggtgtgtgt gctggagaca ggccagctgg	60
ggcctgcage gctctgtcgc tctgtatgt tgacatgggt gggtaactta attatgaccc	120
cagtgcctca agectcggtt tccctgggtt tgagagggag catattggtg ggagggagtg	180
aggactgtcg ggaggggggc tcgttatac aggtcaatctt tggctatgtg ttggctgcaa	240
gggaggacag gcaggagtgt ggaccggaca ctgtcagttt tccaccagg atgaggctgg	300
actgagactg ctgcagcccc gctgggttgc tgggtggggc gagargcagg gggccgggtga	360
ggcagcctca ggaagctgtt gcccctgaaat ccccccgggan gttcccaat ccctctccct	420
ctttgttcaat tggccggggcc tccgtgttggaa aagggttattttttagccctt gcttttttttgc	480
gttaatgggg ttttcggggc ttctctttttaa gaagggttataa wttcaggagc agacatcart	540
gtctgcattt ggactccttg gctcga	566

<210> 34

<211> 1564

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (796)

<223> n equals a,t,g, or c

<400> 34

gaattcggca cgagattttc ctaaccatttgg aaagtttttgc cccaaatgtca cctcgtgg	60
aggatgtatgtttaaatttata aggacttcttgggtccaga catcatgtca agtgcattttgt	120
ctgcattata tttaatcatc acaatattcc taaagggttag ttgtgtatgt tgcatttatt	180
gttttacaaa tggaaaacttgg tggcttagaa agtcaatcgttggttccaa ccctagctgc	240
ccatatttag agtcaattttgg gaacctttaa aacttctcgatggcgaacta caaaaaagcc	300
cgtttaaatc acagtctcttgg aggtgggttgc caggcatcgatggtttttgc agtgcatttcat	360
gtgcggccaa gtttgagaac cagaggttccac acaagggtgcya agtgyagagc tgacatttcca	420
accaggaaag gcccgggtqccaa aqtcacaaatggcaaaaacttgc caccatcagg ttatttactg	480

ggaatttcaa	attatgtgt	caagctactc	tacagatgt	cccttttgtt	ttctcaagtt	540
cttttcatcg	aactgccac	agacttattt	ttctcatcca	ggagtaaaaga	aatggggctc	500
ctgcttcctca	ttacotttgg	gggattctcc	cccactcacg	tttatatctc	ttttaagctc	560
tcatttgcacg	acgtttca	tatatactt	gcaccatggc	atcatctgcc	tagggtttc	720
tgttatttt	cacagagcat	acacatca	gtgtattcta	gaaacagctg	taggctcgta	780
ttaacaaagt	gataanaatc	tggggctwtg	aagacatgt	ggatattagc	taactgaact	840
gtgcttatgag	attgtttccc	tttcttctt	gcaaaccatt	tcagagcagc	ctgggtgtgga	900
gactgtgtt	ctgaaaatggg	gccataggag	cctccagact	cctttctccc	tgccagaggc	960
tgatttaggtt	gatccccatag	tgcagaacts	catgtcacca	ccacatacgt	gtatggcaca	1020
cgtgggccaa	tggaaaagggt	ccaggagttc	ttactttcta	actgggtcatg	tgtaaggaag	1080
agagcacctt	cctcccttaa	caccaagggt	ttcaatagag	accactgagc	tgctacatta	1140
aatgaaaagat	gcattttctt	gtgccctggg	atgctgccc	ttatttcaccc	gtgttttagtt	1200
ttatcttctg	tgtaaaagat	actcgagttt	aaatggctt	aaaatgaagc	aaggggctgg	1260
gcacgggtgt	cacccctgt	atcccagcac	tttggaaaga	tgaggcggtt	ggatcacctg	1320
aggtcgggag	ttcaagacca	gcctggccaa	catggcgaaa	ccccatctct	actaaaaata	1380
caaaaattct	ctatgaatgg	tggcacatgc	ctgtgtctc	aaatactaga	gaggctgaag	1440
caggagaatt	gtttaaacct	gggaggtgaa	ggttgcagt	agccgagatc	atgtactgc	1500
actccagctt	ggcaacaga	gtgagactct	gtctcaaaaa	aaaaaaaaaa	aaaaaaaaaac	1560
tcga						1564

<210> 35
<211> 1035
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (522)
<223> n equals a,t,g, or c

<400> 35						
gcagtttccc	agcaccattc	gttgaagaga	ctgttcttcc	cctgttgaac	gatcttggca	60
gccttgcata	aaaaagatta	accataaaatg	cgagggttca	cctctggact	ctcaagtctg	120
ttccactgt	gtgtatgact	gtctttattt	ttttctatta	ttcttttatg	agattgttat	180
tcaggtgcag	ccacaatagg	agacacttgg	gaggctcagg	aaaaaacaca	gtttatcaca	240
caggtcctag	agacgaggca	tgttgtgcca	tgccatgt	ggccacttgg	gaaagacgct	300
agggtggta	ggaggaggga	argatttagc	tactgcctt	atcgaggtt	ccttagggta	360
gggcaggggc	aacagtttag	tttgaataat	tttggcatac	tttagactgg	cggggtggc	420
tcgttgcctg	gcacctggct	ctgggatgt	aggttagagga	gtgtgcctc	ttggggata	480
agggccagat	agaggagata	tggctcttgg	ttttgaatag	cntgscatat	taaagacata	540
ctcctagctg	ggcccyttgt	tatcttaag	aattggctag	accttggagg	gacagttct	600
caaatacgta	gaaaggttct	ttaacatgtt	aaacatcgt	atacaagaaa	agctaaaaat	660
ccatctgtgc	cagtgcata	ctgtctkgct	tactgtgt	gtgtggtaag	ttttgaaatt	720
aggaagtgt	agccgggcgc	ggtggtctac	gcctgtatc	ccagcactt	gggaggctga	780
ggtgggtgga	tcacgaggtc	aagagatcaa	gaccatctg	gcacacatgg	tggaaacccca	840
tctctactaa	aaataaaaaa	attagcttgc	gtgggtggcag	gcatctacac	tcccgactgc	900
tcgggaggt	gaggcagga	gaatcaactt	aacccaggag	gcggagttt	cagtgagttc	960
gagattgagc	cactacactc	cagcgtggcg	acagagcgag	atccgtctca	aaaaaaaaaa	1020
aaaaaaaaac	tcgtt					1035

<210> 36
<211> 620
<212> DNA
<213> Homo sapiens

<400> 36

acgcaggggg ccaaggggcc tcttcagggc tctgggcacc tccagctgag gggatgtggg	60
ctgagaggag atgagtggct gcaagtggag gaagaacatt acctacttca qattttatgc	120
ctcttgctct gaaacgaggt cagctttcc ttatcccttg gctttcccc cagggagttt	180
gcccgttsga aggtgaacaa cttggctctg gaaaggaagg acttcttcag tttgccatg	240
cctcttgcct cagagttat ccggAACATT cgccctccctg gaaggagacc caatctgcaa	300
caggttacag aaaatctgat taaaaagttac ggcaactcatt tcttaatttc tgccacccctt	360
ggaggttaagc aacatcacaa tcccaagctt attgggttgc agaccattgg aaataaacgtt	420
aagactcggt tagcgttagct ccaacagtct tatcttcytc cgggcctttt attgcctgag	480
aagtctttctt ggtgcatttg aaaacagtgc achtcttca catctcaatt taccctcagaa	540
catctctatt taatttattcc ctgaggaggg gaaagttgggt atggcggtt gagggtggaca	600
ggctcctaaa aaaaactcga	620

<210> 37

<211> 973

<212> DNA

<213> Homo sapiens

<400> 37

ggggacttag tcacacagaa aatagaagaa tttgtgtaca gttggaaagg ctcaagagaaa	60
aggagtttgtt tggacayaat gaccgtttttt tttttttttt tttttttttt tttttttttt	120
caacccacat tacagatgtt acttagtgat agaaaaacatt tccctgtttt ctttcatata	180
ttatgaaata tttactttttt ctagtattttt gtctatctt cgtcaaaagat ttaaatatct	240
ttgacctcct gtactaaata ccacgccaaca tcaagttttttt ttgcctttttt tttttctta	300
ggcttagttt ttggatatacc atttctaaac caatggtagg aacattttaa ggcattttttt	360
gtcttggata wgttttagca tggmccggat gaaagttttt tatgtttttt aattttttttt	420
tataattttt aatgaatatt aatttttttta atgaatataat attaaaccaa ttaataaaaca	480
gtcacaaagc tgcaaccgk ttttataattt attaaagtttt taattttttt atggattttt	540
gtcatctaag ttccgaaatgg aaatacacca aactttttttt tacttttcca aattgtccta	600
ctgtttctca gaatcaacat ttttagacat tatgttagaaa cactttttaa cctagttgts	660
tcaggcttag tagagaaaagg aaaagaaaaga aagttggagc tggaaagagga aagttggtaa	720
atgtggtcag tagtgcattt tttgtgtacca gggcaagttttt gcagaacccctt ttctgaacac	780
cttcacctgtt gtaaaatccc aggcattttttaatctccaa ccactatggc aggatatgca	840
tctgagagca aagaggcaaa tggcaaggcag agatcacaaa ggtgcagag ctagagtagt	900
gatagaacca gtgcaggac gatctaaattt cccttgcattt gtcaatacrc aaaaaaaaaaa	960
aaaaaaaaactt cga	973

<210> 38

<211> 838

<212> DNA

<213> Homo sapiens

<400> 38

cccacgcgtc cgcctccccc tcactctctg ttctgtttt atttcttggg gtttcactag	60
atatctcata taagtggaaat cataaggat tttttttttt gtgattggct tattttcattt	120
agcataatgt cctcaagggtt cattcatgtt ttgtgtatgtt atagaatttc ctttctttttt	180
aagggtgaat aatatttcat tttttttttt gagaggtttt acctttttttt tgattatgtc	240
atgagggttc tgcctctgtt aatgaattttt tttttttttt tttttttttt atgggatgtt gttctgtata	300
aaagaattttt gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	360
catggatgtt tggagcaaga agtcccttca tcaatgggtt gttttttttt tttttttttt tttttttttt	420
ccaaaccttca gaactgtaaa taaaattttttt tttttttttt tttttttttt tttttttttt tttttttttt	480
atagcaacac aaaatggact aaaacagaag atttagagaga tttttttttt tttttttttt tttttttttt	540
tctcaaatgtc caaggtggca taaaattttttt tttttttttt tttttttttt tttttttttt tttttttttt	600
tcatgggttc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	660
gccacaagggtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	720
tattgtttca cagctatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	780
ctgggggaca agatgtcgaga tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	838

<210> 39
<211> 607
<212> DNA
<213> Homo sapiens

<400> 39

tgcacccacg	cgtccgccca	cgcgtccgct	tctgcaagca	caatggtagc	aagaacgtct	60
tcagcacctt	ccgaacccct	gcagtgtgt	tcacgggcat	tgtagctttg	tacatagcct	120
caggcctcac	tggcttcata	ggtcttgagg	ttttagcccc	gttgtcaac	tgtatggtg	180
gactactgtt	aatagcactc	ctcacctggg	gtacatcag	gtattctgg	caatatcg	240
agctggccgg	agctattgtat	tttgggccg	catatgtgtt	ggagcaggct	tcttctcata	300
tcggtaattc	cactcaggcc	actgtgaggg	atgcagttgt	tggaaagacca	tccatggata	360
aaaaagctca	atagcatctt	aacgtgaaga	tcaaacaaga	acacaacaag	cccctactga	420
tttctgggtt	tctgccacgg	ccacaggttc	atatccagag	gaatggcaga	tctgagacga	480
cccaggaaga	gctaaaacat	ggccctgtaa	taaatgagca	gacctcteet	gtggtttcaa	540
attattaaac	acacttccat	ttctcttgg	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaagg	600
cgccgcgc						607

<210> 40
<211> 882
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (198)
<223> n equals a,t,g, or c

<400> 40

gggcgatarg	gtgctgtcct	tgggtgtctg	tgtatatggg	atgatgacgc	ttatcagcat	60
tatctagtcc	tttccacccc	gaaattcgcc	ccgattaaag	actgtgttgc	attatcagcc	120
tgccgaccat	gccccgggc	gtgcccattgt	ccaccttacat	gaaaatgttgc	gcagccagtc	180
tcctggccat	gtgcgcangg	gcagaagtgg	tgcacaggta	ctaccgaccg	gacctgtatga	240
gaaacaggtt	gagaagagtg	aagtgttattt	ctcaaaatgttca	catagcttttgc	gtgagacgtat	300
ttgaggatct	gaagcccaag	ctttctgttt	gcmaaaacttg	gatcataaagt	ctttcggtcg	360
gagaacttgg	agtctggca	gagtcgagca	gaggagacat	gatgactgtcc	tagacttgg	420
ctcagtgtcg	tgttggggag	aactgttacg	acataacctja	aattccacca	aagcgtggag	480
aactcaaaac	ggagcttttg	ggactgaaaag	aaagaaaaaca	caaactctaa	gtttctcaac	540
aggaggaact	taaataacta	tgccaagaat	tctgtgaata	atataagtct	taaatatgtt	600
tttcttaatt	tattgcatca	aactacttgt	ccttaagcact	ttagtcataat	gctaactgca	660
agaggaggtt	ctcagtggat	gtttagccga	tacgttggaa	tttaatttacg	gtttgattga	720
tatttcttga	aaaccgccaa	agcacatatc	atcaaaccat	ttcatgttata	tggtttggaa	780
gatgttttagt	cttgaatata	atgcgaaata	gaatatttgt	aagtctacta	tatgggttgt	840
cttttatttca	tataaattaa	gaaatttattt	aaaaaaaaaa	aa		882

<210> 41
<211> 959
<212> DNA
<213> Homo sapiens

<400> 41

ccacgcgtcc	ggtatttct	aaaacaataaa	atttatagttgc	ttaatattca	tagggtaat	60
aaaaatgaag	cttcccttt	gggcctgcat	tgtatgtgtt	gttttgcaa	ggaagagacg	120
gttcccttc	attggtgagg	atgacaatga	cgatggcac	ccacttcatac	catctctgaa	180

tattccttat	ggcatacgg	aat	taccacc	tcctt	ttat	tatcgcccag	tgaaataca	gt	240
ccccagttac	cctgggaata	ctt	acactg	ca	cacagg	ttt	ccttgc	tatc	300
aacttctcct	ggattccct	atgt	tctatca	cat	ccgtgg	ttt	ccctt	tag	360
gaatgttct	cctccccc	ctag	gggtt	ccc	gttgc	cct	cccttcaa	ggttttt	420
agcagctgca	gcacccgct	cccc	acctat	tg	caagctg	ct	gtgcag	ctgcac	480
tacagccaca	cctgt	tagcag	ctg	agctgc	tgcaagg	cct	gtgcag	ctgagc	540
tggcagagc	cac	ctgttgg	agctt	gagcc	tgtgc	gag	gcac	ctgttgc	600
tgctgcagag	gcac	ctgttgc	gag	tggagcc	agct	gcag	gaac	cttcac	560
tgctacagcc	aag	cctgct	cccc	agaacc	tcac	ccct	ccctcttq	aacagg	720
tca	gtt	gaaat	tct	ttaga	at	act	gtatgc	agaaataa	780
gaaatctaca	aaag	tttttct	tt	cttcca	aagact	att	catt	tgttgc	840
attcatctca	ctacatt	ttgtt	ttgtt	ttgttgc	tagt	tttttcc	ttggact	taa	900
aaaaacat	tg	ataattaa	aaataaa	ata	ataattt	tag	acc	aaaaaaa	959

<210> 42

<211> 875

<212> DNA

<213> Homo sapiens

<400> 42

ggcacgagt	ccagtgtccc	gtgc	ccctcca	gtgt	caaaga	ttt	ggggc	ac	tgcccgtc	60
aatggaaagg	tttgtgtca	gc	cttggag	cct	acac	agg	gggtccc	cag	cta	120
ccatccacgc	accac	ccttcca	ggac	gaga	ac	c	gggtctgaa	ctc	gagg	180
g	cggt	ggaaa	ct	ctcg	acc	tg	gg	ct	gttgg	240
tcggag	ctca	gt	gtc	cggtt	taa	agag	ac	tg	gggg	300
gcatt	ctgt	gcaat	ttcc	gcaat	g	ac	gac	c	cc	360
cac	cc	tttgc	tgt	ggat	gt	tgt	gag	ac	gg	420
ctcc	gttgg	aa	ca	gg	tg	cc	gg	cc	gg	480
aqcagg	ccgc	tga	acaa	aca	tc	cc	gg	cc	gg	540
cc	ct	gt	cc	cc	c	ct	cc	cc	cc	600
gccc	act	ct	gg	taat	gc	at	aa	gg	gg	660
tct	gt	gg	gg	ttt	at	ttt	ttt	gg	gg	720
tat	ttt	at	ttt	ttt	ttt	ttt	ttt	ttt	ttt	780
aagg	aa	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	840
aaaa	aaaa	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	875

<210> 43

<211> 630

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (26)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (29)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (56)

<223> n equals a,t,g, or c

<400> 43
 cccctccctta aggggaccaa agctgnagn tccaccgcgg ggcggccgct tagaantagt 60
 ggaccccccgg gctgcggaaat cggcacgaga aacactgagt tcagctgcc tgcgttagata 120
 ttctggatta ctacagcagc cttcttagcct atggttggaa cccatctgat acttttccca 180
 ttcttgctta gaacaatggt gatatttctc tgcttaaaat cctcttgtgg ctcttttg 240
 cccattaata aaattcaaac tccatttatac cttaatctca tataaaaaac cttcaagatg 300
 tgctccttac ctaactctct tttttccct ttatcttcata tattttcat tttttttctt 360
 acctaaccct gtgccttat ctctttctt gatgatctt ttgttctacc aataaccagc 420
 ttctggaaatg ctatgttttgc ttactcagt cttccatttc ctcttcttgg aattctagcc 480
 tctcttcctt cccctaataatc agtataactcc tactggtcct tcagaactca gtttaggtta 540
 taattctcca aaaaattaca attagttct ctctctggat cccatccctc aaaaaaaaaa 600
 aaaaaaaaaact cgaggggggg gcccgttaacc 630

<210> 44
 <211> 571
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (460)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (494)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (562)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (566)
 <223> n equals a,t,g, or c

<400> 44
 gcaacatcggt ccttactcct gcattttct gacgggcattt ctccccacgca gcttgcgca 60
 caccaagcgg gcggtcaaaac tctcggtca attgttataaa gatcacataaa caaaactttc 120
 atcctggagc tcttcatgc gccgtttgt gctcgctttt ccgtttggcc tgctgcccact 180
 ggctgtcgct catgctcacg aagaccatga ccacgagcac ggcaggctcg gcgcccatga 240
 acatggcgtc gggcgctga atgcccgt ggcacggccag gcccggagc tggaaactgg 300
 cagccccggcc atgaacctgg tgggtttcgat gcatgttagcc accagcgccg ccgacaaaggc 360
 caaggtcgcc gccgtgcgca aacagctggaa aaatccatcg ggcctgttc aacatgc 420
 aagccgcagt tgtgtggtca gcaaccaagg aatcaacagn cggtgttcgt gacaaaccgg 480
 aagccgagca tgangacgat gaccaaggct ccgacgggaa aaggcggcgg cggcccacaa 540
 agcatgatca agaccaaaaat gnaatncacg c 571

<210> 45
 <211> 930
 <212> DNA
 <213> Homo sapiens

<400> 45

ggacaaaatt	gaccat	ttaataatgt	tataaccag	ggctgttac	tttttgc	60	
tttttgc	cst	agtcacaat	tgtaaagca	cgagaaca	cagaatgg	120	
ttcctgg	tag	aagacaagg	acttacg	ggcttctt	atgtggatt	180	
gtt	cacaagg	agatctgt	gctgctta	taattgaa	ttctctgt	240	
at	ttccaagg	agataatctc	cttcttgg	cctaattt	tagatgata	300	
tgat	tctt	gctcattt	agaataactt	tcaggaga	atggcattt	360	
ctt	gggt	cggtataaa	gccaatta	gtacaattt	accataaa	420	
ttt	ctgt	cacagt	ttttttaaa	actgatgt	tttgcattt	480	
aag	tttacag	aaccataa	attctca	tttcttat	tagacatata	540	
caa	actgtat	agatttggg	taaaaagt	tctcagtt	tctccaa	600	
aaaaa	agctt	aattttaca	ttatactt	tttcttaaa	ccatgtaa	660	
cattt	tcaa	cttaagg	tct	gcatagc	ctttaataa	cttgggatt	720
aacaat	atgt	gttctacatt	ttttcataa	ttatata	tgtatgtt	780	
cag	ttt	gtctgt	aaaaa	ctgtctt	caatatgt	aatggtctt	840
tgaa	gtt	tacctgt	tata	taatggat	taaccagt	caataatca	900
ctt	aaaaaaa	aaaaaaa	aaactcg	tag		930	

<210> 46

<211> 437

<212> DNA

<213> Homo sapiens

<400> 46

gttccggac	gccaacatcc	ggccggcg	ggaaagg	gacgtgggt	agagt	gacca	60
tgacgaa	att	agcgcagt	tttgggac	tagcgat	ggc	tttggggccc	120
tgaccac	ggg	agcc	ttgggc	ctggagct	cc	tttgcct	180
tgc	ccct	cttgc	ttgg	cttgc	cc	tttggccac	240
ggc	ctt	ttgc	ccgg	ctatgc	gg	ctgtatcg	300
actt	tc	atgact	gggg	gacgtgc	gc	gagggcc	360
ccg	ttt	gact	gggg	gagct	ca	atcaggagg	420
gac	ccctt	tttgc	tttgc	tttgc	tttgc	tttgc	437
aaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	

<210> 47

<211> 1024

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (5)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (14)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (32)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (713)

<223> n equals a,t,g, or c

```

<400> 47
gtgggntcccc cggntggcca ggattcggca cnngggcgctg gccgccttcc agctgctcaa 60
cctgactggg caacgtgggg ctcttcctgc gctcggatcc cagcatccgt ggcgtgatgc 120
tggccggccc cggtctgggc cagggctggg cttactgcta ccaatgccaa agccaggtgc 180
cgccacgcag cggacactgc tctgcctgcc gcgtctgcatt cctgcgtcgg gaccaccact 240
gccgmctgct gggccgctgc gtgggcttccg gcaactaccg gccttccttg tgccctqctgc 300
ttcatggccgc cggcgtccctg ctccacgttt ctgtgctgct gggccctgca ctgtcggccc 360
tgctgcgagc ccacacgccc ctccacatgg ctgccttcct cctgcttccc tggctcatgt 420
tgctcacagg cagagtgtct ctggcacagt ttgccttgc cttcgtgacg gacacgtgcg 480
tggcgggtgc gctgctgtgc ggggctkggc tgctcttcca tggatgctg ctgctgccc 540
gccagaccac atgggagtgg gctcggggcc agcaactctta tgacctgggt ccctgccaca 600
acctgcaggc agccctgggg ccccgctggg ccctcgcttg gctctggccc ttccctggcct 660
ccccattgcc tgggatggg atcaccttcc agaccacagc agatgtggga canacagcct 720
cctgactcca ggaagagcca gagctgtgca gggaggaagg ggtgagaggg gggcccccac 780
accttagactc agtaaggaag tcgggttggc ccttaacatc tgcatggac aactccaccc 840
cttccttggc cttgccccctg cccgcctaca ctcctacgtg tccagggttt gggccgtgac 900
ttaggcagag gagtgcagag gagggcttgg caggggctgc tcagggccgcc tagctcccc 960
tttgcctaggtaataaaagca ctgacttqtt aaaaaaaaaaaa aaaaaaaaaaaa aaaggcgcc 1020
cgct 1024

```

```
<210> 48
<211> 463
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (14)
<223> n equals a,t,g, or c
```

```
<220>
<221> SITE
<222> (462)
<223> n equals a,t,g, or c
```

```
<220>
<221> SITE
<222> (463)
<223> n equals a.t.g. or c
```

```

<400> 48
gaattcgcca cganttacag gcatgagcca ccgctccgt ccttgactgg tgatTTTcta   60
ctgaaaccca gacatgttag gtattacgag actctgggtc ttactgaaac cttgcttccc 120
acgctgtac tccagcacag gaggggaggt gctgccgcgt tgctgtgagg tggaggccga 180
agtccaggtt ccccactca gacccatgga ctctagagaa gggggcactg tgccttactt 240
tggagggtgtt gggagtccata gattctacta gaccagca ggcctgtgcgg ggtgggggtg 300
tcaccttact gctctttcat ggcctccact ttcaccatat tctcaggatg gcagtgcacct 360
tggagtggag agaggggatg tgattggca ggagacgcag gaggcttcaa aaaaaaaaaaa 420
aaaaaaaaaaa aaaaaaaactc qqqqqqqcccc cqqaaccaat tnn 463

```

<210> 49
<211> 885
<212> DNA
<213> *Homo sapiens*

<220>
<221> SITE
<222> (233)
<223> n equals a,t,g, or c

<400> 49

aattcggcac gagagggctg	cateccttgcgtt	tctgtgagc tctgcccgtt	gggagcatcc	60
atgtctgtatgt gcaggggccc	tgcagcaactg catttttttt	gccttctctg ttctgttttag		120
tacaaccacc ccagcaggc	tccagtttctt gccaggttag	tgtggatggc ccagcaccat		180
cttcctctcca tcttgttggc	tatccctctct tgcgttgcac	aaccccgcca ggntcgccgc		240
tcaggagctc tgccgtgtga	agtgtgtctca gcagttctcc	tcacatgtct acgcaaaatc		300
tctggctccc tgcgtgtctg	agcccaacag acacactgag	cacaggagtt ggctctcagc		360
tcctccca cttgcgtgac	tgagccytgc cgtccgttgg	camgccasg gagaccacag		420
tgcgttgcactg tccaaacctt	acgttaattgg catcccagga	ggagaagcaa gagtgaatgg		480
ggcaggaaaaa gatcattaaa	gaaatcgtgg ctgacataaa	aaaggatgag ttcatgtcct		540
ttgttagggac gcgtggatga	agctggaaac catcattctg	agcaaaactat cgcaaggaca		600
aaaaaccaaa caccatgtgt	tctcaactcat aggtggaat	tgaacaatga gatcaacttgg		660
acacagggtg gggAACATCA	cacaccgggg cctgtcgtgg	ggtgaggggg atggggcagg		720
gatagcatta ggagatatac	ctaattgtaaa tgacgagttt	atgggtgtca gcacaccaac		780
atggcacatg tatacatatg	taacaaacct qcatgttgc	cacatgtacc ccagaactta		840
aaatataata aaaaaaaaaa	aaaaaaaaaa aaaaaaaaaact	cgttag		885

<210> 50

<211> 847
<212> DNA
<213> Homo sapiens

<220>

<221> SITE
<222> (337)
<223> n equals a,t,g, or c

<220>

<221> SITE
<222> (407)
<223> n equals a,t,g, or c

<220>

<221> SITE
<222> (415)
<223> n equals a,t,g, or c

<400> 50

ggcacgagtg aaaccataaa	gaaaaccaag ggggtgataa	atataaaatc caacgtggtt	60
attacctggt tacgggcagg	actatgatag ggaagtcaact	ggttatgttc tgtttcttaa	120
gtctgggggc aggggtacat	gggtgtgcac ttattataa	tgcttccaac cgtataggtt	180
tatTTTatat attctgtttt	acatattgtt gattgtatqa	atgtgtatgt ctcagtaact	240
taagagtaaa tgaactgtga	agaatttagag atggagttt	gagggatttt ttgttttgg	300
ttgttttttgg gagaccagat	cctggcttgc csgtcangt	ggaatgcag tgatgttatac	360
catggcctca cttcaqccctt	taccctctg gggctcaggt	gatcctncca acttnccggc	420
ttcgatttagc tggggactcc	agggtgcacac caccacaccc	agttgacttt taaatttttc	480
gtggagatga gttctccctg	tgttattggcc cacgctggc	tcaaatttcc ggcttatgga	540
atccctccctg agccaggtgc	ctggccagtt ttgggttgt	tttggtttcc ttttttgaga	600
tggcaatttc gctttattgc	ctcaggctgg agtgcagttt	cgcgatctcg gctcaactgca	660
accilucatct cccaggtaca	agcgatttcc ctgcctcagc	tactggggaa gctgaggcag	720
gagaatcaatc tgaacccayg	aggcggaggt tgcagtgc	caagatcacq ccactggact	780
ccagtcctga gcaacagagc	caagactccg tctcaaaaaa	aaaaaaagaaa aaaaaaaaaa	840

aaaaaaaa

847

<210> 51
<211> 580
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (557)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (569)
<223> n equals a,t,g, or c

<400> 51
caagaaaagg ttgcatttata acacaaggtaa cacatgttaa gtcctttt acaatctctg 60
tttagttgcac tcaatgttct ttttctcctc ccaaacttct tagcaattt taaaaacac 120
acctacgatt gttatTTTtag gttcttccca attttttttc tgccctccaa ggaaatgtgg 180
ta:tctctgac ttccacagct ctcaacttga ttccataacac caaggcttct tccccctgga 240
gtgggagaca caatagggat ggggatgtgg ccagaagaag aggatctat gaatatccat 300
agttttatcc accccacccc aaatgttattt atattaaagg acccactgaa gagcgctgt 360
gaqatggtag actcttgtct gaagtcctta tgcacacca atataatctt cttcttaggc 420
cttctcccac taggtgaaag gggataagct tgggacatc tagaaagggg cattaatttc 480
cccaatccca tgagctacat ttgagactca cagtattaga aagccggggc tttatcacag 540
tctctttgga agaagcnaaa tttttttcng gacagcttgc 580

<210> 52
<211> 598
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (515)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (523)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (558)
<223> n equals a,t,g, or c

<400> 52
ggcacgaggg tcggcgggca gggcgggct gcatgcgatg acgtcgatgt gtccggccgc 60
ggcgctggcc tggccgacct cggcgatttc ctgtatcgtc agcctggcgc ccagctggc 120
cgccggcgcc gacaactggg cccgcgtcgcc atacaccacg caggccgcgc cagccctgcg 180
cgcagccctg acgacgattt cggggccgat gcccggcggca tcgcccattgg tgcggataat 240
gggcaggggag gggttcacgg tgctggaat gggattgcgt tgcggataat gctacacccg 300
ttcgcgccgc acggccgggc ttactgcgttgcgat gacgcccgc 360

tcgaagggtga ccgtggcgca cttgggagcc caagccggtg gcgttgttgc gatgtatctc	420
cggcttgaag cagcttaacgt tcccatggat ttgcataatgc tgcttcggc gcttccaacq	480
cctgggttct tcagcttcca acccaagtcc agcancttgg ctccccggaa gcttttaagt	540
ttaacctggt aaagtgtnc cttgtcgaa cttggcttgc aacccatggc ttcaatgc	598

<210> 53
<211> 571
<212> DNA
<213> Homo sapiens

<400> 53	
gaattcggca cgagtcccac cagccccca aaaaacctct cagtagtttc tttcagtgt	60
aaaaatgatg agcatttttc tatgtatggg ttttaaccat tattcagggt ggtctttgt	120
ttttaaatct ttttttaact aataagatc acgggtgtgt ttttatacag aatgcatta	180
taaatgttt taattgtgtt ctgtttttc cagtctttaa gtccatgcc aattgttctt	240
atattctata qaagttcgct caaaataactc aacaggggaa taggcagcgg acagtcagaa	300
tggttggaat tttggcttc taagaaaaac ttatatttc ataagcatgt ggtcagatca	360
ttttgtcat atgcagcctg gattggatgt taagtaaatg cttgttcagt gccggtacat	420
ttacttaat ctgttttat ttttgcatac tagaataacta ctgtgtcat cataatgtaa	480
tctatattcg racctttttt tttttttttt actttgaagt cttaaataaa alyialalaata	540
cccaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaa a	571

<210> 54
<211> 1247
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (2)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (9)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1131)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1202)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1209)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1226)
<223> n equals a,t,g, or c

<400> 54
cnccccacanc cacaattgg ttgggttggt tgtaagtca gcttcccacc aaagtccaaat 60
atttctaaca ctctagtgc aaaaaatta ttattaaata gctaagagg gtgcattgtgg 120
gaaaggtcag tgcataatccc tttaggagg gagaatgtg taatataatca gctatcgagt 180
tgtttaaaaa aagtgtatyc aaycgtatat tgtctatagt atgtgctatg aaatttgcatt 240
ttatgatatg taacaggggc aaagccaaat tcattgttact ctgttcagtc agaaacattt 300
tgtggcatac agcatccctg ggaagtgcg tactttgtt cgtttgggtt ttagtttgc 360
attnagatgt ctttataatt gatgcctatt ttaatagcat ttcttttag cttttgggtc 420
gtatccat tcactgttcg tatctgttac ttcttattaa agcattatct gtttaccaca 480
tgtacaaaaa ctcttgaat aatatgcatt octagtttc agccaagtcg gggatgttag 540
tgattgtacc agcccmaagc acttggataa tcagggccct tttccctttt ataatcaatc 600
atcaacatca gaaaaagcta cttgtttat ttatattccc ttccaaatcc gctctggAAC 660
atgcagtaac tgccaaaaac ttatTTTGT aacaaatatc attggcaact ttggaatata 720
tttgatattc cattaggatt ttctaaaag gggaaataaa ctatataat atatgtatct 780
taccccaat tcttccaaca gaatttctat aggaagccat ggatgtatggc ataagttgc 840
cacatattac atgattttaa ataattctca aaatacccaa ggaactctta aagagtttg 900
gtatgagtat actactttgg tttaattttt gtttcatgga tggttgcatt ggaaggattt 960
ttgtttcca cattttcca ttgttagcag agtggaaatcc aagagaccaa acatttgcaa 1020
gcattgtatt tgacacttt tgtaaaaaaa aaaaaaaaaa aaaaaaagga aaatataat 1080
aataacttaaa aaaaaggatct taggaaggg ctaccctca gattggggac ntctcttaac 1140
cctaccccg ggaccccggg ggagggtatgt tggccctatg tgggggtctg ttattccat 1200
tnnnnnntt ttttagggttq qtaatcnnnnn tgggggggtt tttttccc 1247

<210> 55
<211> 848
<212> DNA
<213> *Homo sapiens*

```
<220>
<221> SITE
<222> (8)
<223> n equals a,t,q, or c
```

```
<220>
<221> SITE
<222> (15)
<223> n equals a,t,q, or c
```

```
<220>
<221> SITE
<222> (98)
<223> n equals a,t,q, or c
```

<400> 55	
cccccaanga acagntttg gtaaatccaa aaattatgcc actttaaaaaaa aaaaaaaaaaca	60
aaccttaaaa caggatctt aataccctcc cttgttnic ccctgcttt actcctccac	120
tttagaatat ttcccttaaa aatcacctca aaggactgtg aggaaaggct gtggtacctg	180
accttgttga aatcaaggcc cggcactgta ctacaggcct gtttacagat tattacggtg	240
aactgaatgg gtaccgagge ttcaccaaag aggtactttt ttgttgttgc tggtgttta	300
ggaataattt taccatttt aagagcattc cccccacaccc tccccacacaca cccaaacaaa	360
atgtgggtt gttgccttca aaaaagagaa gttttgtgtc attaacatga cagaagaact	420
ttttaaaaaa aaataactgt caactattct atttgcatatt aggagactgt tmatctatgc	480
tagattgtca tttccctcc ttctcccaca gaagtttact ggtagtccat gtcatggctc	540
gtagctatcc ctctaaccat accatggaaa tgcaggcacc caatgtgaaa aggagcaatt	600
gctgggcattt actgacaccg ctcatgtttt acacatgtt gagtaatcag catatctaga	660
attatcttgc attqccttaaa tcataatgtat atagtgaatg ttatataata tacctggcag	720

gtctgtttta atttaaattga aaaaaagatac aaataactttg tttggctggc aatattaaa 781
 ttattatatg aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaagg 841
 cggcccgct 841

<210> 56
 <211> 669
 <212> DNA
 <213> Homo sapiens

<400> 56
 cagcctcatt ttctcagtgc cccagaggtc taggatagga tttctaaact ggaatcatcc 60
 ttaatcacct tgaagatccc ttaagaggca tttgactggc gctgcgtct gtgtcctcaa 120
 agcaatgcgt gtggcatcggt cctgtgtaca catgcagac taatacccaa actaaaaact 180
 gggtaactgg ccctgaagtg ctccccatc agtaagccac agggaaatgt ttgattttta 240
 tgttctgtt gattttgggt tggttgcac atctaaaggt gctttactt ttctttttt 300
 ttttttttct ttctgtttt ttttgttagga cttgttctaa catggaaaac aagtccagaa 360
 gactctccctc tgactgttac ctttgcucca agccacccaa aacttttatg ctcatgtttt 420
 attaaagcag gtgccttcgt gaatctctgg gacattttg aggcatttga agcagaatat 480
 aqagtggctt catctccctt cttaatcttc ctgggtgggtt ggatgttcca cttgtatcat 540
 aqattttttt attacagata tgctccactg tttttaaatg tgaacctgtg cgcaaattgtg 600
 cagattcaat gttcttgtta cagattqaat aaatttttat tttgaarawr aaaaaaaaaaa 660
 aaactcgag 669

<210> 57
 <211> 680
 <212> DNA
 <213> Homo sapiens

<400> 57
 gttccatgtg gcactgacta ctccagagtc cttggaaaca ttagtcttt tttgcacact 60
 cgtctccatt ttccacgcta ttagtctttt atatgcaagg tactcttagt agttgcagca 120
 tctgtccacc gtccttggct ccgttagtata accgggtgtc tctttactaa atgagagtat 180
 tcctcagtttggagtcga gagagagaag agggagacag agagagagag agagttgggg 240
 tgcttatttt aactctggtt ctaatcatgc tgcaggtcag cttgccaaga aacacttaag 300
 gctctgtttt ctgcattgttag gcaggtaatc ctctgataga gaaacataga gttatcccat 360
 caaaatgtga gcacagaaat gtatcaacaa catgccataa gccagtcgt atatctaaaa 420
 gccaaactgtt aaaccggtcc ctgtccccct gaaagggtgt cagaatataa ttgttttgc 480
 aggttttaat attaattggc actagacaca cccagctgaa gcaaggctcg tcgttggtt 540
 tgagtttcct ctgccttttta ctgatgttac aatgttttaac tttggcttttgc ttagatttt 600
 atgcagacca agagtgttat ttttatgtat aaaatttttat atttataaaaa caaaaagtac 660
 cttaaaaaaaaa aaaaaaaaaaa 680

<210> 58
 <211> 524
 <212> DNA
 <213> Homo sapiens

<400> 58
 gggccatat tctccggatg agtgcatctc tttgcctgtt tacacaagtg ctgaaaggga 60
 tcgtgtgggtt accaatattt atgttccatg tgggggcaac caagaccagt ggattcagt 120
 tggagcagct ctattccataa aaaatcgtttaa gaatctaattg acaacaaaaa ccatcttac 180
 aaaaggaaac attgattttttaa atagctttaa atcaaacatg tggctagtct acatttggaa 240
 tggtagttca aaatattaac atatagttat gtgttgcatttgc tcaactgaaat ttatgtgt 300
 aaaaggcagca ctgtgcattt tttaaagttaa taaattaaatg gagttattgt taaaacagag 360
 tattcttttgc acaacattaa atatttctgtt gagaaggatc acttttccatg tggctcaaaa 420

atttgttta ggtccaga ttttaagtgg tatattaacc aataaaat attttggtcg 480
 tcaaaaaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaggcgcc 524

<210> 59
 <211> 427
 <212> DNA
 <213> Homo sapiens

<400> 59
 ggcacgaggt catttcagcc ttatgaattt cccagaataa gctagatcac cttaaggcc 60
 atgtggtag ggaaacttgg gcacagaatt tacatttca acttgggtat aagatgggtt 120
 taaggttaga atcaaataagg agaaagcctt agctgttcca gcggcccatg ttAAAAGAA 180
 tggcttctt tttccaagta tttctgcgc tgcatgcac tgagttctt tggAAAGGAG 240
 caccatgcag gcatatttc cagacaggac cggatttgct cgttactcag aggtgtgtgc 300
 attcttgc ttttaggatat ttaatttaca tctttataa gtgatattac ggtgtcttaa 360
 aagtttatgc atttggaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa 420
 aaaaaaaaaaaa 427

<210> 60
 <211> 1263
 <212> DNA
 <213> Homo sapiens

<400> 60
 ggcacgagcc cttccggca aaacggcagc agtagcagag gcagcttctg agagcctggg 60
 caggcagcag ctggctgacc aagtccactg gaagagaagg cttgtgccag ccgggagaag 120
 gaagccgggg acaggatgaa agcaacaaca ccttgcaga cagtcgaccg gccaaggac 180
 tggtaacaaga cgtatttaa gcaaatttac atggtgcaca agccggatga tgacacagac 240
 atgtataata ctccttatac atacaatgca ggtctgtaca acccacccta cagtgtcag 300
 tcacaccctg ctgcaaagac ccaaaccctac agacctctt ccaaagccca ctccgacaac 360
 agccccaaatg ctttaagga tgcgtccctc ccagtgcctc ccccacatgt tccacctcca 420
 gtcccgccgc ttgcaccaag agatggctc tcaacagaaa agcatgactg ggatccctcca 480
 gacagaaaag tggacacaaag aaaatttgg tctgagccaa ggagtattt tgaatatgaa 540
 cctggcaagt catcaatttct tcaagcatgaa agaccagttt ccaagccctca agtttctgaa 600
 ttgattgtca ttctaatggc tcaagattat atttcaattttt gaggatttt tccaaattatt 660
 taattatatg caaacaaaaaa tatctataca tttaaagaaga accacttgc tcctcgaagc 720
 gcagggtcac acataggtct tcccttggcc ttctttgggt tttggattcc ccaataactcc 780
 ctgcattgtc ctttgcattcc cctcagaaat cagtgcctc aggattttct tggaaactcc 840
 acagttccctc tgcatttcct tccaaagccctt cttcttaggaa aacatttggaa atcaattgca 900
 atttactctc aatttttggt ttAAAAGAAAT ttttaggtgtc agtgaacagg aaaaagagag 960
 ttttctttgt ggagctccctt caagagtctt caagaaatca cactcatatt aatccaaggc 1020
 cagccctgtg gtttattgg aaactcttagt cttgccacac ccaacatgag tgccttgc 1080
 tgagtctttg aaaccctgtg tttgcgggtt tttcttggag aagcatgttc aaaaccttat 1140
 catcagataa ttcatcataatg atatttgc tttttttttt accctttttt taaataactc 1200
 atatctttta gagtaacagg actacaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa 1260
 aaa

<210> 61
 <211> 720
 <212> DNA
 <213> Homo sapiens

<400> 61
 ggcacgagat ttctcacaat gacaaattct caaatattgc taatagtact gtggattttc 60
 ctacatttgtt aaattgaagg aattgctaaa tgctgaattc agcaaccagt ttgagattgt 120

tgaaaataaa gattgtttct tccaaatgc aagttcacag atcactggag	tagctac	180
agtttggct agaccagagg ttgcagatat ttttgtccca taaagagaca	catggtaat	240
atttttggct ttgttagttg tatagttttc gtgttagctg ttcagctctg	ctacatgaag	300
caaccataga ccatacctt acaagtggc acctttgagt accaataaaa	ctttatttag	360
aaataacaga gggctggatt tggctcaqt ttgctgaacc ctttctaga	tgaaggctcc	420
tcttgccaag actggctccc taccttggct gacaaattct cacttggga	cttagtcatt	480
gttgcgtc tcgttattt tgcatgtctt ttctcatgtt taggtgctgt	gtcttaatac	540
ttttttctta cattaattt aacaatcatt actgagcgct ggtatgtcta	gtttcttttc	600
tcttctttcc tcctttctt ttctttttt cttttctt atttgaaggc tctcactctg	tcactccagc	660
ctgggtggca gaccaggacc ctgtctctaa aaaaaaaaaa aaaaaaaaaa	aaaaaaaaaa	720

<210> 62

<211> 589

<212> DNA

<213> Homo sapiens

<400> 62

ggcacgagcc ccgcgcatgg ggaggttaggc tggaccggc ccgcggagct	gctgcagtcc	60
tgcgcgcct cctgcgcctc cccaccacca tcatgtccca gttcctgctt ggatttacac	120	
tgggcaacgt qtttqdaatq tatctggctc agaactatda tataccaaac ctggctaaaaa	180	
aacttgaaga aattaaaaag gacttggatg ccaagaagaa acccccagt gcatgagact	240	
gcctccagca ctgccttcag gatatactga ttctactgct cttgagggcc tcgttacta	300	
tctgaaccaa aagctttgt ttctgtctcc agcctcagca cttctttctt ttgcttagacc	360	
ctgtgtttt tgcttaaag caagcaaat gggccccaa tttgagaact acccgacatt	420	
tccaacatc tcacctcttc ccataatccc tttccaactg catggaggt tctaagactg	480	
gaattatgtt gctagattag taaacatgac tttaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	540	
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa ggcggccgc	589	

<210> 63

<211> 686

<212> DNA

<213> Homo sapiens

<400> 63

tcgacccacg cgtccgaact gccaaaagct ggtgattctg ggacaggcct tcactttgga	60
gccacggat ggggtgggg agccccatgg gcctgggaag gaggggtgctg tggagggggc	120
tgcagggctg accagcagge agcctcatct ggtcgggggc gggggcggca ggacgagaag	180
cggggctccc gtccttggga ctgtcttgggt tgccacggg ccctgaggat gcacgggtgcc	240
tggggctccc gtgcgggtgg gcggggggca tgctggcctc tgagcgatca ggcgaggcca	300
gcgagggtgt gcttgc当地 tcaagcaata agaggggggt tcctggggc ttccagccca	360
ggctagaagc ccccatggct tctggcagct ggacatcagc cccaggatt ggggtgatt	420
tggtcatgac agtgtgcctg tcccactgtt acacgcata atgggggtta tgggggtgggg	480
gtgggactca aggcttgacc gactcctagt ggacctgatg taaaattctt gtcaaaca	540
caccacttt caatggtttg cttaggat ttcgttattt aaagttctta attatgtttt	600
ttaaaaaaaaaat actaaaaata aagggttcaag ctgc当地 aaaaaaaaaa aaaaaaaaaa	660
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	686

<210> 64

<211> 452

<212> DNA

<213> Homo sapiens

<400> 64

attgtgttagc attcaaaacc tgctgaacac tcaccctcaa tgttatattct ctgttctggc	60
ctgcttcaag gtcagttaca ctacttcctg ggatggcct ttctctggct taagcttgg	120

tgtccctggc ttccaaagg atcacagccc	aaaaggcaca gtggaaaaaa tttatggcct	180
attcgagaag agtgagcagc ctaatcaagc	caagccttga tttggggttc tcacttcact	240
ggtattttcc ctctgtcccc aattgggttt	tctatagttt ctagtttcc aggccctccaa	300
gggagattct gaggcttgat gtgttctgac	tgtgtcttgg ctttgtatg ctgagtgcca	360
gaaatactct gtactataaa aactaccatc	gttctttgaa acaacaaaaga ggaataaaga	420
acttaattct ggtaaaaaaaaaaaaaaa aa		452

<210> 65
<211> 370
<212> DNA
<213> *Homo sapiens*

```
<400> 65
ggcacgagtt gcggtgaacc agaattataa cagttagctc atctgactgt ttttagatgt      60
acagccctagt gttaaacattt ttggtatctt tttgtgcctt atctaaaaca tttctcgatc     120
actggtttca gatgttcatt tattatattt ttttcaaaga ttcaagatgtt ggcttttgtc     180
atccactatt gtatgtttt tttcatttgac ctcttagtgat accttgcattt ttcccacttt    240
ctgttttcgg attggagaag atgtacctt tttgtcaact ttacttttta tcagatgatc     300
aactcacgta ttggatctt tattttttt ctcaaataaa tatttaaggt taaaaaaaaa     360
aaaaaaaaaa 370
```

<210> 66
<211> 987
<212> DNA
<213> *Homo sapiens*

<210> 67
<211> 1018
<212> DNA
<213> *Homo sapiens*

```
<220>
<221> SITE
<222> (1014)
<223> n equals a.t.q. or c
```

<220>
<221> SITE
<222> (1015)
<223> n equals a,t,g, or c

<400> 67

ggtgcggcccg	ggggcacggc	gctggggagg	cgtatggcgcc	ggccgcgtcc	aggctcgccc	60
ccgaagccgg	gctcggggcg	ctgccgcggc	ggcgctcgc	ccagtacttg	ctcttcttc	120
ggctctaccc	ggtgcgtacc	aaggcggcca	ccagtggcat	tttgtcagca	cttggaaact	180
tctctggcca	gatgatttag	aagaagcgg	aaaaagaaaa	ctctagaagt	ctggatgtcg	240
gtggggctct	gagatatgcc	gtttacgggt	ttttcttcac	agggccgctg	agtcaacttc	300
tctacttctt	catgaacat	tggatccctc	ctgaggtccc	cctggcaggg	ctcaggaggc	360
ttctctctgaa	ccgcctcgtc	tttgcacccgg	ctttcctcat	gttgtcttc	ctcatcatga	420
actttctgaa	ggggaaagac	gcctcagect	tggccgccaa	gatgaggggg	ggtttctggc	480
cgggcgctgag	gatgaactgg	cggtgttgg	cggcactaca	gttcatcaac	atcaactacg	540
tccctctgaa	gttccgggtg	ctcttcgcca	acctggcgc	tctgttctgg	tatgcctacc	600
ttggctcttctt	ggggaaqtga	cgaccgctgg	gagaacatca	gtgtcaactgt	ggacgtgggt	660
ctgggggtct	caccgcggca	gcgagagcag	aaccaatcca	gtcaggatgt	cactgactct	720
aaatcagggt	attcaagatg	ccccaaaatg	atggatagag	aaacagaaaat	ctgtgaatgt	780
cagaaccctg	tctttaaaaa	aggcagtcrc	tgccttcagg	tggtgcgtcc	ccagaaaactt	840
aaaatttagt	cgaggcagtt	tcaattgtta	ctgtggacgg	aattaggat	acaataaaacg	900
ataatgcagg	tcttcaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	960
aaaaaaaaaa	actcgagggg	gggcctgtac	ccaatcgccc	tgtatgatgt	ctgnncac	1018

<210> 68

<211> 762

<212> DNA

<213> Homo sapiens

<400> 68

ggcacgagtg	cctctacgtc	atgtcttctc	caatgtcatg	attcacgtcg	tgcagtaactg	60
ttttggacct	gtctattatg	tccttgggg	cctaactgtg	ctgagccaaag	tgcataatgg	120
tggcaggaat	gcctacataa	cagggaaaaaa	tctattgtat	caagcacgg	gttccatata	180
tcttgggatg	atgtatgtca	tctggtcate	tcccccata	tataagtgc	catgttattc	240
gcggcaatct	cagggaaaat	aaagcaggag	tggtcatca	ctgtacccac	aggatccat	300
tgggagactg	gtttgaatat	gtttcttccc	ctaactactt	agcaagagct	atgtatctacg	360
tttccatgac	cgtcacctt	gggttccaca	acttaacttg	gtggctagtg	gtgacaaatg	420
tcttctttaa	tcagccccctg	tctgccttcc	tcagccacca	attctacaaa	agcaaaattt	480
tctcttaccc	gaagcatagg	aaagcttcc	taccatttt	gttttaagtt	aacctcagtc	540
atgaagaatg	caaaccaggt	gtatgtttca	atgcctaaagg	acagtgaatg	ctggagccca	600
aagtacagg	tcagcaaagc	tgtttgaaac	tctccatcc	atttctatac	cccacaagg	660
tctactgaat	gagcatggca	gtgccactca	agaaaatgaa	tctccaaagt	atcttcaaag	720
aataaataact	aatggcagaa	aaaaaaaaaa	aaaaaaaaaa	aa		762

<210> 69

<211> 630

<212> DNA

<213> Homo sapiens

<400> 69

tcgacccacg	cgtccgcctg	gttctcaccc	ctgttggagt	cagctttcaa	gattgcctgt	60
ggccctgctc	tgccccatcc	tggctggcag	ggctgtcatgt	ttccatctcg	tttgcctctt	120
tatattaatg	ccaaagtttt	agccaaagac	atcttcctac	ttttgttgt	tttcagcatt	180
ctgtttctgca	ctgtgggctg	gctctctgc	ccaaaccctgg	gcactggccc	ctggctgggc	240
catttcatgg	ctcaaaggct	ctggggctc	aaggaaggct	gggctgtca	gtcccttc	300
gggtcttgc	aatgaaaatg	agcatatatg	tgtttaaaaa	atattaatcc	ttttgaaaag	360

aactgagaag araaaatgtat aatttttatcc catttttaat attttagtct agcaacttgt	420
gatacataga tgacaatttt gtgagtttt caaatgtgt tacagatttt tgtaaatatg	480
actctttgt aattaactca tgtacagcct catcctgtat agtttaatga tgaatgtgca	540
ggggacctgt ctcaggctcc tatatgtta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	600
aaaaaaaaaggg cggccgctct agaggatccc	630

<210> 70
<211> 940
<212> DNA
<213> Homo sapiens

<400> 70	
ggcacgaggg gactctgggc atggctgtgt ccctgtggcc tgaggggtct gggcctctct	60
gtgccctctc cctgctcacy tgctgcctcg ttttgcgtcc tgcttcatcc tctggattcc	120
tctggagcct ggaagagact cctgccttc aaggggcttt tgaaattgcc cagcctaag	180
tgttagagtct agttttaga ctaattttt ttgtatTTT gtatttgtaa tttttgtggg	240
tacatggta acatgtttgt aggtatgggg tacgtgagat attttgatat aggcatgcag	300
tgtgtataaa tcgcattcagg gtaaaaggggg atccaaaccct caaacatTTG tcctttgtgt	360
tccaacaatc caatttaatas tcatagtgat ttaagaatgt atgatacatt attggtgctt	420
gtaatcctgt tgcgtggca aacaccaggc cttaactgggt cttttcact atgatttgta	480
cgtgttcaacc ctccccaccc ctgttagtct agtttttattt ctgtcttttc cccagaggcc	540
acagaaaactg gtgtggagag gtgtatgtca ctcccagcc ttggggtcag cccctgcatt	600
ttctctctag gaccatgtgt ttctcgattt ttttcattcc tctgtgtca gtttagccag	660
atgcagtcat atgagggtct tgcgcaaaccc ttagggtaga tacagatctc aaaaatgtgc	720
caaggggggg gagggcagga ggcacaggaa ggaggagagaa ggaaggcttag ctgactcctc	780
ccaccccaa gggctgcccc tccctacact cctgttttag taagagtggc agccacctgc	840
cctgtccccg ggggtgacaa ctttgagctg ggagggcag cccctcagtg cttttttta	900
gacctgtctc acctgcaggc gtctctctag ccaggctcaa	940

<210> 71
<211> 1103
<212> DNA
<213> Homo sapiens

<400> 71	
gtcttaatga gcaacagcaa cagcagtctc cagtttaagaa agagagaatt aaatacagca	60
gagatttcct gttgaagctc tcaagtgttt ccatctgcag aaaaaaaccg gactttctgc	120
ctgatcatcc cattgtactg caaaaaccag aaaacaaacca aagtttaag tagcatttt	180
agaacagatg aatttaagt tggacatctg caaatgaggt ggatcttagca acaataactg	240
taatggactg tgacaattca atttattctt aattttgtatg gttggctatt tgacttctct	300
aaaaatgaga aagagctatt taaaatata aagaattttc taatcagttt cagctttgca	360
ggaggtttcc tgcataatg gggaaagtaac actggaaagt aggaatttgg ttagtgaagt	420
ggaaagactg tatattata atttgcatac tacttgcaat tttttgtttt tcatcacttg	480
taataatgga atggaaatgt aagctgtaaa gactctcaaa tataaaatat ttgctacagt	540
gtatatatgg tacataattt cttgttgctt taaaagttcc ttctgttgc ttgcttccca	600
ctgatttcat accagctcat gaatggatca ttacagtctc tccagaggct tagaatgatt	660
cagaatgttc aatgcatagt tctcaataaa caggaggcag aatttttaat gggtatTTCT	720
tttcagatat atgattggc tctaggtttt tgataataat atggctttaa attcataatt	780
actagcagag attgataatt tggaaacaat ggttagtgaat gaaactgaag ttgaaaaacg	840
gctgctactt atgtcactaa tcagaccata tgaatagcag aagttgagca atttcaaagt	900
aaaactgata ttttatttc caaaggaaatt tagacatttg aaaataattg acatacatta	960
agtttaatt cgataatttcc ttatatatgg atgaacaattttt tttgggtta agtttaat	1020
tcctagaaaat ttatacatt aaatctctg caatttgta ctctggatgt tactgtttaa	1080
aaaaaaaaaaa aaaaaactcg tag	1103

<210> 72
<211> 899
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (20)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (85)
<223> n equals a,t,g, or c

<400> 72

ggcagagtgg	ggagggcgtn	ggtggatgag	tcaccatsct	ccctcaagga	cgtcttggc	60
aaggtgtctg	gcccccattag	ccagnaaacca	ggaaaatgta	gctgcaggaa	aatcacctcg	120
tttccctcggt	atgtttttc	ttaggctggt	ttcccttaca	agctgcaatt	atgttccatc	180
ccacgcaatt	cagtaagtg	gcactttca	gagaaactgt	cttggtgata	atttgggctg	240
ctgtqqqac	aaaaqtqagg	agagaaggja	gtggagagrt	caactgagrt	ragrrggttt	300
tgtgtccatg	agccattttac	aaactttqca	cctgattggg	ctcagttgca	gtttcttgta	360
tttcccttacc	agccaagctg	ttgaagctqc	tgagccccga	atgatgttat	cactgaggca	420
gatgacaaac	cctcttagttg	ctagaaaacca	aactgctccc	cgagctggtg	tttccgtttt	480
ttgtactgac	tgcttatttg	ggcttgatat	ataatggta	aaacaggaac	tgtttatttt	540
agggtataag	aaaccaacat	tatgacaaga	agatgtcattc	ttagttaccc	tgttaccagt	600
accataggcc	agataactatc	taqatgctta	taaacatctt	atctaattct	tgtataataa	660
agccccaaagc	taggttttat	atccccattt	tatggatggg	ggaactgagg	ccaataactt	720
catataactt	atccaaggcc	acaaaactag	taataaacag	agtgaaatcc	aacccaaaaa	780
caaactacaa	atccaaattt	ctttacctct	atgctgtctg	tacttgctgt	tactagcaaa	840
gttcttttgg	tgggagttac	cccatccccct	ctccaaaaaaaa	aaaaaaaaaa	aactcgtag	899

<210> 73
<211> 549
<212> DNA
<213> Homo sapiens

<400> 73

tcgacccacg	cgtccgggct	gacatgatgt	atttctgccca	gatgctggca	gttgtggaaa	60
ctatcaatgc	agcaatttgg	gtcactacgt	caccgggtct	gccttctctg	atccagcttc	120
ttggaagaaa	ttttatttt	tttatcatct	ttggcaccat	ggaagaaatg	cagaacaaag	180
ctgtggttt	ctttgtgttt	tatttgtgga	gtgcaattga	aattttcagg	tactctttct	240
acatgctgac	gtgcattgac	atggatttgg	aggtgctcac	atggcttsgt	tacactctgt	300
gattccctt	atatccactg	gggatgtttt	gccaagctg	tctcaagtgt	tcaagtccatt	360
ccaatattca	atgagacccgg	acgattcagt	ttcacattgc	catatccagt	gaaaatcaaa	420
tttagatttt	ctttttttct	tcaattttat	cttataatga	tatttttagg	tttatacata	480
aattttcg	acctttataa	acagcgcaga	ggcgctatg	gacaaaaaaaaa	aaaaaaaaaa	540
aaaaaaaaaa						549

<210> 74
<211> 590
<212> DNA
<213> Homo sapiens

<400> 74

attggatcgt	tttccctactg	ggacgtggcc	ccagcttctc	cgtgactctg	cagcacacct	60
------------	-------------	------------	------------	------------	------------	----

gttcccaccc	tgttc	ccc	caggattgtg	ctggaaagtgc	tgg	cgct	ccgaagcata	120		
agcgaacagt	gcc	ccgtgt	gtcc	aggccag	gtc	accgttg	cctc	agagct	gagacacagg	130
cagtgggtgg	aaaggacgt	gcgg	tcgc	cagcggcaga	act	acctgcg	tatgtggagt	240		
agtatcagac	tactgtcccc	tgt	gctc	agc	ctgata	actgt	tactcattgc	gctggagt	300	
gtcaacatc	atgctgtttg	tgg	gaagaat	gcgc	atgat	atc	agcgtta	cctaaagt	360	
gtaaaagtgc	tcttgc	agta	cacggagaac	ctgg	tggt	acacc	atgtt	cgaaaagaac	420	
aagtggaaatg	aaactatcaa	tctt	acrcat	acag	ctt	tgaaaat	tg	gacttt	480	
gagaagaaaac	aatgttaat	acat	tttagcc	aaga	aaatcc	caag	taaagt	act	tttatga	540
aaacttgtaa	aaaaaaaaraa	ararr	raaaaa	aaaaam	ctcg	aggg	gggg	ggcc		590

<210> 75
<211> 1056
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1051)
<223> n equals a,t,g, or c

<400> 75										
ggcacgaggc	gaaaccgcct	acgacgg	tcgc	actggag	ttcc	aaagagc	ctct	gagttc		60
ctgc	c	ttt	tcat	ctct	ga	atcccc	acca	ttgg	ccctaca	120
at	gat	gt	gt	ac	gg	ggatt	tc	gg	ggatgtgg	130
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	240
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	300
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	360
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	420
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	480
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	540
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	600
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	660
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	720
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	780
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	840
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	900
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	960
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	1020
tt	tt	tt	tt	tt	tt	tttt	cc	cc	ccatag	1056

<210> 76
<211> 930
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (919)
<223> n equals a,t,g, or c

<400> 76										
gcagagctgg	acag	cg	cctg	ctg	ccc	ccgc	ccct	ccgc	ccatgg	60
gagccacttg	gc	ct	cc	tcc	cc	cc	cc	cc	ccatgg	120
g	ct	ct	cc	tcc	cc	cc	cc	cc	ccatgg	180
g	ct	ct	cc	tcc	cc	cc	cc	cc	ccatgg	240
g	ct	ct	cc	tcc	cc	cc	cc	cc	ccatgg	300

caaggccagc atcccaggc c	ttcgat gggaccctta tcacctctcc a	gacctgt	360
actacctcac tgtggagtcc tcagagaagg agagctcag gacccccaaa gtgggttgcaca			420
tccccactta cgaggaagcc gtgagcttcc cagtggccga ggggccccca acaccacctg			480
cataccctac ggaggaagcc ctggagccaa gtggatcag ggatgccctg ctcagcaccc			540
agccgcctg gcctccaccc agctatgaga gcatcagect tgctcttgat gccgtttctg			600
cgagagacgac accgagtgcc acacgctct gtcaggcct ggttcagact gcacggggag			660
gaagttaaag gtccttagca ggtcctgaat ccagagacaa aaatgccgtg ctttctccag			720
agtcttatgc agtgccctgg acacagttagg cactcagcaa acgttcgttg ttagaggctg			780
ttctatttat ctattgctgt ataacaacc accccagaat tttagtggctt aaaataaatac			840
ccattttatt atgtcaaaaa aaaaaaaaaaa aaacttcgta gggggggctc cggtacccaa			900
tcgcccctgat gagtgagtna tattgtttccg			930

<210> 77

<211> 4463

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (3308)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (3469)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (4119)

<223> n equals a,t,g, or c

<400> 77

cagcaatgaa atcctgcttt cttttcctca gaactactat attcagtggc taaatggctc			60
cctgattcat gttttgtgga atcttgccttc cttttttcc aacctttgtt tatttgtatt			120
gatgcccttt gccttttctt ttcttggaaatc agaaggcttt gctggcctga aaaaggaaat			180
ccgagccccc attttagaga ctttggtcat gctttttttt cttgcgttac tcatttttgg			240
gatagtgtgg gtagcttcag cactcatgaa caacgatgcc gcaagcatgg aatctttata			300
tgtatctctgg gagtttctatc tacccttattt atattctgtt atatcattga tgggatgttt			360
gttacttctc ttgtgtacac cagttggcct ttctcgatg ttcacagtga tgggtcagtt			420
gctagtgaag ccaacaattt ttgaagacct ggatgaacaa atttatatca ttacctttaga			480
ggaagaagca ctccagagac gactaaatgg gctgtctca tcgggtgaat acaacataat			540
ggagttggaa caagaaccttggaaa gactttaag acaaaaattttag atccttggag			600
ttctttttctt gtgcttcagt ctccctgtctg gcacctttgtt gcacagactc cagctgacat			660
agttttttttt gattccctt tcatgcttc aactcaagggg atgagctggg ctcagcttgg			720
gttctcttctt cctgcacatc ggcctggaaa ctctcaagac aagaggcgaa aaaaggctt			780
agcatgggaa agaaatttgg tttttttttt ctctttttttt ttgagacatc			840
catctcggtc ctcttgggtgg ctgtatatat tctttgccta ttgggttgcgtt aaacagcaat			900
gcacaaaagga acaagggggc ctggaaatagg aaatgcctctt ctttcttacgt ttggttttgt			960
gggagctgcg ctgtttttttt ctatctttagt gtgtccctgtt ttgtcggttt			1020
ctatagcctt cgattttttggaaa gaaacttttac tcccaagaaaa gatgacacaa ctatgacaaa			1080
gatcatttggaa aattttttttt ccatcttgggtt tttttttttt gctctgcctgt tgatgtcgag			1140
aacactgggg ctttccatccaa ttcacttacc aaatacttca agggattttag aacacggccaa			1200
gccttctgtt aatgggcatttccaa agaaaggactt gttttttttt gttttttttt gttttttttt			1260
caagagaccy ayadacilccag atttccacgaca ttccctgtccc atgttagaaggc atttccattt			1320
awccgtggsc cctcttcaga accttagaccc attcgttccaa ttttttttttt attaattttttt			1380
aagaacttgg ctatggctga tttttttttttt attaactttt ctgtatggacc ctgttagttt			1440

cagtttaagt	cagat	ctt	acagacatat	agaacagcgc	attc	tgt	agacatttg	gc	1500
tcatgttgt	aaatacaatc	accatata	aaaaattgtt	ttcacctgat	atgaaaatgt				1560
tagaaaaggc	aaactccggg	acttctaaag	atttacttaa	atcccattat	gtactttatt				1620
cagaatgtag	aagctgactt	gaaaggcatc	cttggacta	agtgaagctt	attcagaaaa				1680
tgcattttc	aatgcata	gcaactgctt	gtagatata	ttttgcagt	gtatgttga				1740
gctgtatgg	ttgcaattat	gtttcttatt	tcctaaaag	caaaaagcgt	agtttctgat				1800
ttatgttata	gaatgata	gattagactt	tgagccaagg	ggaaaatact	aaattcttt				1860
aaacctggag	ccttagagag	ccacaggaat	atcttctgtt	gtacagtcta	ataagctgt				1920
gttaggaagta	tcatgtatc	acagtttaat	gacagtttat	gtataatat	aattcagtat				1980
tccctctgat	aacatagtt	ccagtgttta	atacacttgt	aacttggatt	tttacctt				2040
aggctatatg	tatactcagt	tttttaaagc	attttttca	gagatcactt	aattccccat				2100
gcttctgcaa	tgcataaaa	aactataat	gccgagtgg	agaaaactcct	ctttcttcat				2160
agtccctcagg	cttggttac	atttgcata	gccat	gcctccagct	tttaccagtt				2220
taacatccaa	agttcacagc	atcagcattc	atgggttaa	aacagtttt	cagtataaca				2280
cgatctgata	atcattcagt	tattaaattt	taaataatta	ttgggatgg	ttcttggctt				2340
taagtccact	gaataaaaac	tatgaaattt	cactctgt	caaccatcca	ctaggataga				2400
ataccgaaat	ctgtgcata	aaaaatagga	gatgggccc	tttgcacaca	attcgtat				2460
atgcagtctg	ctatataat	atgtcacat	gcactgtgt	tatgaaaata	gatggtctgt				2520
gttcagacaa	aagtaaaaaca	ttttttcaa	attgttacat	ttaaagg	tctggagaa				2580
atttatgaaa	cgcaggctgt	gtctatttga	catcagaaat	ttccactt	aacc				2640
ataagaaact	ttaatctgt	tat	taca	tttgcgtt	gag tacacttccc	ccttattt			2700
acgtctgcat	ttccctccga	gttcacatc	tttctaaat	gcagcttgg	tttaaaataa				2760
aagaacattc	atttgtgt	tctaaacaag	tttcagtaaa	taccacc	atagta	ctgg			2820
tgaatttctc	agcataaaat	cgacatac	aaaaagtt	taaaattt	cg	ctctttcca			2880
atttcattt	tatgcattt	gaagtattt	ttgcagg	ttttag	tgaagcttgg				2940
agtccatact	ttgagcagac	caagtgaagg	gaagaacaga	aagaaactca	ggagtaga				3000
aatatcactt	ctgcacttac	accacttca	ggcacatcca	aagat	ttcct	agatacttgg			3060
aaaatgtctg	aaaat	tttta	agtaaatac	taaactt	ttc	agtgtttagc	tcaactt		3120
gttcatttgg	aagtttctct	ccatccgagg	acttaagcc	gttttggatt	tgtaagcc	c			3180
gagtacaata	cacttcttgg	aggcatctc	actgtgtt	aagcaaa	gga	tatgc	at		3240
gtggaggac	ggcttgcac	cttggact	ca	tgcc	ttt	gaacaaat	at		3300
ccttgggntg	ctgcgtatc	acggttc	cttgc	gg	ttt	cttgc	at		3360
ctgcctctgg	agagactgac	tccaaat	tgca	gttgc	ttt	ggact	gt		3420
gctttatatg	acgaactccc	agaaatggat	gcc	cc	ttt	ccat	at		3480
gcycctggta	aggacagtcg	ctctgggg	cctcattt	cct	ttt	gcagtt	tc		3540
agtgaaagag	aggagataga	ccctgg	agg	cag	ttt	actgc	tca		3600
ttctggagct	acagttt	ttctgact	gg	atagg	ttt	gat	tc		3660
gcagtcc	tattctct	cctttagt	at	ttt	aa	ggcatt	actat		3720
ccagagttt	tttggggaaa	cttgcagatt	ttt	attt	aa	tttgc	aa		3780
aattat	tttta	caattataa	ttt	ttt	ttt	tttgc	ttt		3840
ccaagata	tcttagattt	actatttac	tttgc	tttgc	tttgc	tttgc	tttgc		3900
acctctaaat	atgaggctt	acatgaaat	tttgc	tttgc	tttgc	tttgc	tttgc		3960
ttaaatttgc	gcactaagg	gttgcgtt	tttgc	tttgc	tttgc	tttgc	tttgc		4020
ctcagatgt	ttaaaggcgc	acgtgtac	tttgc	tttgc	tttgc	tttgc	tttgc		4080
gtctgggt	catggcc	cctgc	tttgc	tttgc	tttgc	tttgc	tttgc		4140
agg	tttcat	catttacagg	tttgc	tttgc	tttgc	tttgc	tttgc		4200
cataaaggca	gaatttctac	aagt	ttt	tttgc	tttgc	tttgc	tttgc		4260
ttttttaact	tgaaaaaaact	ggtataat	tttgc	tttgc	tttgc	tttgc	tttgc		4320
agtaatttgc	ttttgttac	agatt	tttgc	tttgc	tttgc	tttgc	tttgc		4380
agataactt	agtgtactt	ttttaat	tttgc	tttgc	tttgc	tttgc	tttgc		4440
ttcaaaaaaa	aaaaaaaaaa	aaa	tttgc	tttgc	tttgc	tttgc	tttgc		4463

<210> 78
<211> 791
<212> DNA
<213> Homo sapiens

<400> 78

cccacgcgtc	cggttctcct	gcttgccatc	aatggagtga	cagagtgttt	cacatttgct	60
gccatgagca	aagaggaggt	cgacaggtac	aattttgtga	tgctggccct	gtccttcctca	120
tccctgggt	tatcctatct	cttgaccctgt	tggtgtggca	gcgtgggctt	catcttggcc	180
aactgctta	acatgggcat	tcggatcacg	cagagcctt	gcttcatacca	ccgctactac	240
cgaagagccc	ccacaggccc	ctggctggcc	tgcacctatc	gccagtccctg	ctcgggacat	300
ttggccctcag	tggtggggtt	actgctgttt	cgaggttattt	cctctgtgt	gagcagggct	360
ggccagccag	actggcacac	attgctgtgg	gggccttctg	tctgggagca	actctcgaaa	420
cagcattct	cacagagacc	aagctgatcc	atttcctcag	gactcagttt	ggtgtgcccc	480
gacgcaactga	aaaaatgacg	tgacttcagg	gaagecttgg	cacccgaggg	acctggacca	540
gctatggta	gttctgtggg	tggAACACAT	tctgtgtaa	agcccactg	agggctctgc	600
agcggagtga	cagcaacccc	agagatgagg	caccagagag	tgccactgca	tgagacacct	660
gtgaccattc	gaagtctgaa	atgcgggggg	ggagtttcat	tttaagtga	agaccaaaaa	720
cccttaaaa	ataatagttt	tttatcattt	tatagaaaaa	aaaaaaaaaa	aaaaaaaaaa	780
aggcgcccg	c					791

<210> 79

<211> 1292

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (488)

<223> n equals a,t,g, or c

<400> 79

tgcacccacg	cgtccgacta	ctttataat	cagaaaaatag	gatatataga	aatggtctgg	60
aaaaaatgag	gagagaagtt	acaataggga	ggtcataatcc	aatctgtttt	gaccagcaag	120
taaaagcagg	aaatgtgtc	caccaccaat	ggcttaaata	tgtgtgttgg	atgggtgggg	180
tgggtgggtt	ttctggggtt	ggggatgggg	ggAACCTTGG	gatgtgtatgg	ttttcttagt	240
cagagatggt	gttttacagc	tgggaagtat	cttgaacttg	gtggaggtct	atccagacat	300
caggcagatt	tcatattttc	acagacaaag	gctacgttta	cgtctaaggg	aaaaacacaa	360
aatactaaga	tagaaacctc	catgccccct	caccttttca	gacaacaaga	accccccaggg	420
cagagggtct	tcctcactct	cagagttact	ttgacttctc	atctggttc	gtgtggtaa	480
tgcatttnca	ctcttgagtt	tctttttttt	ttaaggatatt	ttcattatga	cgtttggttct	540
tttgcaaaga	gctaattctg	cagaattcta	cccaggggagg	gccgagggaa	tttagtgaag	600
gtaacgatac	cagttagaaa	gttggatttt	ggtcatactca	ggggcctcct	ttggcctct	660
ttggctttagt	ttcttagagat	ggcccttact	tcccaggaat	agatagtgt	gtcttactga	720
gttgcaggtt	gagtctgtat	tgaagcgcaa	gattttcttc	cctgacttgt	ctgggaatcc	780
ttacccagga	ctgtggcttt	cctttccat	atccctccac	caggargcag	gctctaagcc	840
tctgtgtgt	catcctttt	ccacttccat	gaacatgcct	tcaaatactt	aaaaaaagtca	900
ctatacgcta	aagtatactt	ttctttttt	gtgatgtttt	tctgcatttt	ctctataacc	960
agagagtaag	aataaaaacta	cagcttggga	cctggcgccgg	tggctcacac	ctgtaatccc	1020
agcactttgg	aaggctgagg	agggtggatc	acctgagggtc	gggagtttga	gaccagccctg	1080
accaacatgg	agaaaaccccg	tctctactaa	aagtacaaaa	ttagctggca	ttgtggcaca	1140
tgccctgttgt	ctgagctact	cgggaggctg	aggcaggaga	atcgcttcaa	ccttgggggt	1200
ggagggttgca	gtgagtttag	atcttgccat	tgactccag	cctggcaac	aagagtgaaa	1260
cggccatctca	aaaaaaaaaa	aaaagggcgg	cc			1292

<210> 80

<211> 1283

<212> DNA

<213> Homo sapiens

<220>

<221> SITE
<222> (341)
<223> n equals a,t,g, or c

<400> 80

gaattcaaag tttaccaaat gtgcaaaatg agcagttta ccttaggatt attatTTTA	60
tttatattta ctactgcaga aaattatTTG attCTTTTC agagaaaATA ctgtttggTT	120
atatTTTGGG gggagTTTg aatttcacat acgaaAGAAA taacacAGCC CTTCaaACT	180
gcctgtgtt caacctgcaa agTTTTTTT gtgctaaAGA tttgagCTT gtGAaggATT	240
ccctttttgt tccttcttcc ccagcaatct cagctacCTG ggCGCTCCTG ctaatgATT	300
ctggggTTCC gtGCCAGGGG tcggcaggAC aagtGTTCA ntGAGCTT catttggTT	360
ggagtCTCTT CCTCYTCTGA GCCwacAAAG CTCGGGTCCA CGGGTACTCT GSCAAAAATC	420
atcatCTTAG TtaggcATTt ggcagaATAG gtGAGGCAGG gatGAATCTT taacAAATGT	480
taatGTTGCT ttGCTGGAA tGTCAGAGG ggCATCCAAG atGAGCACAC atttAAAAGT	540
aaACACATGA ataAGTGGCA gtagAAATTt ttttGCAACT ctGAGTGTCA cAGTGTCTAC	600
tGAATTCAgT gtATTCCACG ttCTTATTAC aactAAAGAC tGGGTAGAAC GGAATTCTCT	660
taactATGCA aaggAAAAT ccaAGACAAG attCCGCAgg ctGCTGGTGA aaAGGGGTGT	720
tatcatGCAg atGTCATCTT AACAGATTAG cAGAGGGAAg tGGAATGTT CGAGGATGTT	780
caatGCCMCg ttGTTGGTTW trGCAAAMCC actGGAAACa MCACAGGAGt CTAAAATAG	840
aggCCTGGTA gggAAAATGG tacAGCTACG gaATGCAATA CTATTGAAgC ATTAGAAMCA	900
atGAGCTCT GACAGCCCCA gagAGTTATT cataATGTGT agTTAATTt AAAAAGAAAAG	960
tcGAGAGTCa gACTCTACAA gggcATAATA CGCCATTTG GtAAAGAAAAA tGtGtATGTA	1020
gatATGTAaA tagATTGGA tacGAATTAT tGtATATACG aAGGAAGAGt GCCAAAGCCT	1080
acataACCACG CTTTAATAG ttttaATCTC tGTTTATTAA AGAAAGATTG AGGGAGATGG	1140
gatttCTGTT ttTATTTAT acaaATCTGc attGTTGAA tttttttt ttttacGACA	1200
agCTGTTATT tCTCTGGGA GTTAAAGAAA AATACAAAAA AAAGGGAAATT CGATATCAAG	1260
cttATCGATA CCgtcgacCT CGA	1283

<210> 81

<211> 708

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (40)

<223> n equals a,t,g, or c

<400> 81

aggCCAGGCC tccACAGACC CCCATGGCCC CCAGGGACGN aggggaggAC agAGCCCTC	60
agaACAGAGG CCTCATCTCA CTGCAtcccc catCACCCCCC tagTTCCCCA atGGTCTCAA	120
tttGTTTCT GAGATCCCAG tttACTCCGT gGCCAGGGCC cacCTGTGTT tCCAAGTCGG	180
gCTGGAGACG CAGGATGGGG tagGCCTTGT GtCTGAGCA ACCCCAGCTC TGCCtCACAG	240
gcAGGCAGGC CCGGTGCAAG AGTGGACTCT ggTTCCtAA AGCAATAAA GCAAACAAAGC	300
caACAGCTCT GtGCCTAGC aATTtCCATC ttagCCACAC TTtCCCTTC AGGGGCTTCG	360
gaggAGAGGT CAGGGCTAAG GCCGGGGATG AGACTGCAGG AGAGAGAGCA GCGGAGGGCC	420
acATTcGGAG CCTCCGTCCA CTCCAGTTT ATCAGCTTT GCTTTGCA CGGAGTGCTA	480
aACAAATTCT AGCTCTGTGTT tttttcccA ttcccAGATT tactATCAGT tCTCCTTAA	540
AAGTATCTAA GtGTTACAG tagCTTCCC ttCACTTGTAT tCTATTGTGT GTTTCTATG	600
tttGGAATAA ttACACCCAA ATATCTAGAT ATTTCTCTT CACCGCATT tGtAAATAAA	660
GAGATGTGA tgCCAAAGAAA AAAAAAAAGG GCGGCCGc	708

<210> 82

<211> 1464

<212> DNA

<213> Homo sapiens

<220>
<221> SITE
<222> (15)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (63)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (132)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (887)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (889)
<223> n equals a,t,g, or c

<400> 82

gactgtgc	tt gcagnaaagag taaacactct	cggatgccgc	tgtcctgggg	gagcccgcgg	60
gangctgtga	atgttatac gagctggcca	gtccctgggc	cagctcaatt	gtccagctac	120
ctgccagg	tgnttcactg tgttaaaat acattgcatt	ccaagctgg	cccctctgtg	180	
tatca	ctgagaaaatc ctgccttagt	tgttttggg	tgtgtcctag	catttacaag	240
aaaatgaaaa	gcgtcccttt aattggcacc	cgaatgttgc	tgtggctcag	tcacatatcc	300
aaaa	cggtgcgtcccc	cgagccccga	gcccctctgc	agctcaccc	360
ttggcttgtt	tccgcaaacc	cggttaaacgc	aagcccttgg	ggcagatgca	420
agggaggg	aacctgcctc	tgggtcaccc	tgttagcaca	gcgttctcat	480
catggaaactc	tctctcgca	tgctcgaggc	tgtgtgtcag	tgtttgtctgg	540
ccttttttgg	ctggataaaag	aagtgcgtgt	ttttgcgtgt	cttctgtggc	600
cctcacggat	tgaccggag	atgagtgcgc	atgaccacgt	ttaaaggag	660
cctgggtggg	ccctcggggt	ggtctcaggt	cacatttgc	gtctgcaaca	720
cccggtccag	agcgtggta	gctttgttttgc	ccttctgggt	cggtttcgc	780
gtgtgtgtta	aatccagag	cccagaggaa	gtgcaagcgg	gtcctccgccc	840
gcctcttcgc	ggcgctgttg	gacamagc	gctgtgaatt	cgcgtanang	900
tgaaaacacct	tcctgagtag	tccggccttgc	tcaatgagtg	gggagttgtt	960
ctgacatatt	tactcgta	tttcaaaacca	gaagcatgaa	aggaggaga	1020
tccgtttaac	tcgatagaaa	gcgcaggggg	atggcccccg	gcmcaggctc	1080
cagcgctgac	cccaccggcc	tggccggaggc	acttggcctt	ttgacccgct	1140
cctcccttc	atgaccgggg	tgaattagaa	cgtttttaaa	gacacccct	1200
gtaacacatt	gtaattggag	aagaaggaaa	ctctgcaagg	tccaaattct	1260
tggctacaca	tagactctag	tcagtttgt	ctccagaacc	ttaggcttt	1320
attttaattt	cactgttaat	ccttattgtc	tttttattt	agatgttgg	1380
gtagttgtgc	ctcaattatt	gcaaaaatgt	aacaataaag	aaagcaggag	1440
aaaaaaaaaa	aaaaaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	1464

<210> 83
<211> 616
<212> DNA
<213> Homo sapiens

<400> 83

tcgacccacg	cgtccggtca	aaatagtgg	aaatttagtag	aaataacata	ttttatatcc	60
aatttagtct	ccaaaatccc	aacatgcact	cttctgtata	cgttttcag	tatgcttgac	120
tggAACGCC	aattctacag	tagtctgg	agcaacatac	tgcagattaa	ataccttagt	180
agcctatgtt	cttgaatgcg	gacataaagg	agcaatgctt	ttcctatctt	aaaaaaacag	240
tttatatgaa	tgaaacttct	gttctgtta	agatattata	tgttggttag	tgttagttgc	300
aaagcaacta	gcacgattcc	aagtaatata	gaaatcacca	gcttgagttg	ggtctgccat	360
aacagcacct	aaaacgtatc	cactaaatta	gtattaaatg	gacaagtaaa	ccaaactcag	420
agggttggaa	tgaagacttg	taataccag	tgaaaaaaaaaa	ttattgaaac	taccatctaa	480
aattaattgg	aagcttaata	ttacctctag	gaaagagtgt	gggaaatgag	gaaagggcaa	540
aaggtaatgt	gttccagttt	gttctgttcc	ataatcccag	gaaatagata	aacaccaggc	600
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	616

<210> 84

<211> 928

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (916)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (917)

<223> n equals a,t,g, or c

<400> 84

aaaacaaaaaag	gagacgaagg	acgcacgt	ttgggtgagtc	ccggattctg	gtgggttctt	60
ccgctcaggc	tgggtgaagc	gcttccgggt	cgccgccccgc	agcagccstcc	cggcgcgatg	120
aagacactga	ggctcagaga	ggttaagtga	ctcagccaag	gtcaaacagc	tagtaagtgg	180
tggagccagg	actcaaagcc	agtctaggag	ccatgtccac	tttggcccc	tcactcttcc	240
ctcgtgtgac	tgagactctg	tggtttaatc	tggatcgacc	ctgtgtggaa	gagacagagc	300
tgcagcagca	ggaacagcag	catcaggcct	ggctccaaag	catcgccgag	aaagacaaca	360
acctgggttcc	tattggcaag	ccagcctcag	agcactatga	tgacgaggaa	gaagaggatg	420
atgaagatga	tgaggatagt	gaagaggact	cagaggatga	tgaggatatg	caggacatgg	480
acgagatgaa	tgactacaat	gagtcaccgg	atgatggaga	ggtcaatgag	gtggacatgg	540
aaggcaacga	acaggatcag	gaccagtgg	tgtatctaggt	agacaaggca	gggtggccctc	600
agggagattc	caggccagcc	caaactaccc	tgcatcccaa	cccccaaccc	ctgcccacag	660
aaccagctga	tggccccagt	gcctgaaagt	gcccttgggc	acctcctcag	ctgctgccag	720
gatctggct	ctttggggccc	tcccaggcca	tcagtctgca	ttgaaatcc	ccagggccctg	780
aaacctaactc	cacccctctg	gccagtaacct	caccccttga	ttgccaggtc	tggcttaagt	840
ttcttaata	aagacaaagg	agtgtatccc	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaactc	900
gggggggggccc	cggaannaat	ttcccccca				928

<210> 85

<211> 723

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (722)

<223> n equals a,t,g, or c

<400> 85
tattgttagtt agaaccatct gacacatagc ttttattcca ttggttttt gttatgtctt 60
tctttacaag aatttgaagt ccatcaggcc gggagtttg tttgttgtgt ttgctgctat 120
ctccccagtgc ctaaaattgc ctggcataca gtggcattt aataatctt gaatcagtgaa 180
aaaccagatg gtggcttggc atttccacat aggaatgagc caggtggaaa tcattccagga 240
tataagtaga tcttgaagtg ataaggaagg gtcatacataa tcattgtggg cccatttgc 300
cctttttgtt ttctttctc taggctcagc aacagcctca ccaaggactc catgaatatac 360
aaagcccata tccacatgtt gctagaggtg agagcagctc accccactac cagactctgt 420
gttttaggggtg gtgacactgaa gaaggaagag agcgaaagaa gggaaaggacc atttttccct 480
ctaaactgga gtcaagggag ggaggtcaga gcaaggctgg gggcgttaacc cagacccagt 540
ctttgttcaa tctcttctgt ccttttttc aggggcttag agaactacaa ggcctgcaga 600
atttcccaga gaaggctcac cattgacttc ttccccccat cctcagacat taaagagcct 660
aatgccttt gaaaaaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa 720
ang 723

<210> 86
<211> 570
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (6)
<223> n equals a,t,g, or c

<400> 86
gaattncgca cgagctcggt ccgtttcatt ctgttttagaa gttcatgttc atcttagcca 60
tttggaaacctt ctcatcttc tatctttctt ccacgggtggc tgggcttgc tgcaaatcat 120
tgtgtcaaaa tcaaactatt ttcaaaacag cccttgcctt ctgagccctt cccctaagtc 180
ctctgtgggg gtccatgatt ctgcagaggt atgggacaga atcttcagat tttacccctt 240
gagtctcttc ctatgtcatat cctggttccc tcattctaattt attgacaaag gatgactcat 300
taagtgcac tggttatgta actttcaaat actttcattt tgtagtgcag gatctgagga 360
acaaaatgat gtcatttaat cgaaatctaa atgtgacaca aacaacgtgc cagcaataacc 420
tgcttgcgaa ataatgttct gagccacag tggcctggg tatgtgagtt tatataaagt 480
gaaaaggctg cttaatttgc attaaagttt tggaaatgtaa agcttcaaaa aaaaaaaaaaaa 540
aaaaaaaaaaa aaaaaaaaaaaa aaactcgtag 570

<210> 87
<211> 639
<212> DNA
<213> Homo sapiens

<400> 87
aaaaaaaaatgc tagggagaca aaatcaaattt ttaaggggtt gggctctcag cacatttttgc 60
gtttgcattt tccagttgggt cagaaggctg acaatccgc tagcctctgc tttgagcgtc 120
aggggaccctt gttctattcc tgcatttttgc gccatcatctt acacactttt tatctttctt 180
tttaaatttt taaaattttt gaaatctata tacatataag ccataatgttc aacttaaaga 240
atagtaaaca actgtgtccc taggatccaa gtaagaaat agatcagagt cagtttcttgc 300
gaagcttcttgc tatgtgcattt tccccagttca tggctctcc tggctctacc tgagggaaat 360
tacagatttc atgctttctt ttatagttt cctttacaca cataccctta agcctctaag 420
tactatatgg ttccgggtttt gaaagcccag aaggcttattt taatgtgttataaagaata 480
tgcttagccgg gtatggtgac tcatacctgtt aatcccagca ctttcagagg ctgtggcagg 540
agggttgctg aaggcttagga attcaagacc agcctggca atataaggag acccccttcac 600
tacaaaataaa aaaattaaaa aaaaaaaaaaaa agggcggcc 639

<210> 88
<211> 708
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (14)
<223> n equals a,t,g, or c

<400> 88

tacggacacg aggncaaaaa tgagaaaggtaacaatttcg aaaaagcatg cccttctgct	60
gtgtttccag ttgttagat gtctgcctc catgtatata tggatcacat tcgttgttaga	120
tggaaagtgt ggaatccact gttctctcaa accggctctt ttcccttgta cctatcatag	180
tgtacatagc tcaacttcct gagtttgatt ctatgttca aagataggta tttttcatat	240
aagatgtcct gtcaaagcaa gtcatttgaac ttacctggta tttaactgaa aacaaacaaa	300
aatcagcaat ctcttccatt gctttagaa atactgactt aggccaggca cagtggctca	360
cgtctaattcc cagactttg agaggccaag gcaggagttt cattttagcc caggagttcg	420
agaccagcct ggcaacatag tgagaccttgc tctctgtaaa aaggaaggaa ggaagggaag	480
gaggggagggg tggagccaga ggaggggagg ggacactctg ttataacttat cgaaagggtgc	540
tatccaggtg tggtagtgca gccgatagtc tcagctactc aggaggctga ggtgggagga	600
tcacttgagc tcaggagttt gaggctgcag tgagctatga tggtaaccatg tactccagcc	660
tggcaacag agacagacca gactccctaaa aaaaaaaaaa aaaaaaaaaa	708

<210> 89
<211> 949
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (55)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (508)
<223> n equals a,t,g, or c

<400> 89

catgataacc ccaactcgaa taacccac taaagggaca aaagctggag ctccnccgcg	60
gtgccgcgc tctagaatta gtggatcccc cgggctgcag gaattcgsca cgaggttgt	120
tgtgtgtgtt gcttgggtgt ttgtctctt ttaatggtc gtgggtgaca agtgtgtcag	180
agtacttgc cctcctatat gtgtatctat ggcacgtat ctttcttgc gtgtctgt	240
ctgtatgtt gtctcttctt agcgagtggc tgcaggtatg tggccctcg ggggtttct	300
tctgggtcca tggatgtatgactatctgtt ctcttgc tttccctgtg tagcttttag	360
tgtggtttctt ggatttttc tttgcaacga tagtaagcgt actctgcatt cctgtgcatt	420
gtgtttgtgt gcaggtataat gcttccctat tatgtttctt ttctgacttg atttgcgt	480
agctgtgtgt gtacacggct gtgtgcancat atttgctgaa atgcgttgt gtgtgtgt	540
gtgtgtgaga gagagagaga gaggagagag agagagaagg agactatggc ttttctgttt	600
gkmcaaarrt catgtsagcc tatgagtgc tctctctgtg actggagctg tatgtggta	660
catgtggta caagtgcaca ttcaagttca catacacaga gatatcattt tagggcttga	720
acctggaaagt ttgcctccag ggtcatctga acctggattc aggttcagat ccagggccat	780
ctgaacctgg atcgtgtgtg tggaaagac ccaggaccac cacacaatgt cakcagctgt	840
gtgttaattgt gtgtctgtg tggctgtg aatctgtgtg tgtgatttcg ctgttgatttgc	900
tcttggcat ggctgtgggt ccacggccgg tgagttcag gagtctcga	949

<210> 90
<211> 1171
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (291)
<223> n equals a,t,g, or c

<400> 90

gaattcggca	cgaggctctc	cagtgcctat	gtaaaagaaat	gaaagtattc	aggaaaagct	60
gtgcgttgtt	ggaggaccgc	gttgtggatt	tggccacctt	gtcacttgaa	gtcgtgttag	120
gggtgcgttc	gcaggatgtg	gccggatgcc	tttcatgatg	atgctgcaca	gmaaaactgtt	180
gtctttstg	gaagctttgt	ggtaactacgg	tggggggct	ttccttgct	gtgccggctc	240
tgtacctact	gactgttatt	ttggggggct	ggaccaaaga	agacttgtca	ntgataaaatg	300
tactgagaag	agcacaggac	tcctttaagt	ctcaagggtgc	tctgggctta	gttcttctga	360
gcaggggaaac	cagaggctgg	cgytctgtt	tcttktgtta	aatggaaaaa	atacctgcac	420
ttqccactta	actaagtac	tgaagagatc	atgtgcattgg	aagatgtaaa	acagttatggc	480
tctttataag	taaggtggca	ttattacitg	agctggtgga	aggcagcactg	tttcccacaa	540
ttggtctcaa	aagcccggga	tgccctgtga	gttgcattt	agtttattac	cttagcaaaag	600
cagagttggg	ggtgcgattt	tcgatagtag	gctttggag	aatgatkt	tatattyctg	660
aataaatgt	gtccttgaga	aactcataag	ttgcaatgtt	atcctgtctt	aattgtgttg	720
ggcacractc	ccactgcaat	accttaaata	actgaaaaca	tttgccttgc	aaagccccaa	780
tcgacttggta	caataaaaac	agttgcattt	tttgccttag	agatattttt	tgccgtttcc	840
atcattccac	tgcctgggta	ttccttaggg	gaataacaga	taggatactg	gggcttcacc	900
actattttagt	caggatcatg	tttgaatata	agaatctctg	ccttatgtt	atagtaatcc	960
ctgttagtttag	catgaaaaca	aattgccagt	ttgattttct	aggacagctc	aagcagaatt	1020
tgtaccacta	ggctgttaat	tttaagtatc	taattttctt	atttgaaaat	gtatgattta	1080
aaaattggaa	aaagtttttgc	ttataagctt	caaaaggatt	tactataatt	acaatacgtt	1140
aaattacaaa	aaaaaaaaaa	aaaaactcgat	g			1171

<210> 91
<211> 1151
<212> DNA
<213> Homo sapiens

<400> 91

ggcacgagtg	tcaatgaaag	tgtttctaatt	gcaactgcga	ttgactccca	gatagtctaga	60
agtttgcaca	tcccactcac	ccaggatata	gttgtgtacc	caagctatga	aatttagcaaa	120
cagagactca	gtattgtcat	tggcgtgtt	gttggcatta	tgacggtgat	tctaatcatc	180
ttaattttagt	tgtggcaag	gtactgcagg	tccaaaaata	aaaatggcta	tgaagccggc	240
aaaaaaagatc	acgaagactt	ttttcacaccc	caacagcatg	acaaatctaa	aaagcctaaa	300
aaggacaaga	aaaacaaaaaa	atctaagcag	cctctctaca	gcagcattgt	cactgtggag	360
gtttcttaagc	caaatggaca	gaggtatgt	agtgtcaatg	agaagctgtc	agacagccca	420
agcatggggc	gatacaggc	cgtaatgtt	gggccccggca	gtcctgaccc	ggcaaggcat	480
tacaaatcta	gttccccatt	gcctactgtt	cagtttcatc	cccagtcacc	aactgcagga	540
aaaaaaacacc	aggccgtaca	agatctacca	ccagccaaca	catttgtgg	agcaggagac	600
aacatttcaa	ttggatcaga	tcactgctt	gagtagct	gtcaaacc	taacaagtac	660
agcaaacaga	tgcgtctaca	tccatacatt	actgtgtttt	gctgaattcc	actctaataat	720
gatgtccat	tatgcaccat	actgtgtatg	cctttctact	ccgaaacctg	ctggagccctg	780
cccttggccg	tgggtgtca	gccaatca	gcttgttcca	cttgttgc	attttatttt	840
tgagtctt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	900
cagtattat	gcagaaatgt	gtctactatg	gatgtctgag	tcaccagaaa	ttccatttctt	960
aaagaggcgg	ttagcaccta	cagtgtatgt	tttttttttt	tttttttttt	tttttttttt	1020

tattgcaaca ataag[REDACTED]ga gactttgtgt gaacaaaggg aaat[REDACTED]gcc tcttatgtct	1080
ttgtcttaa tacattaaat actgattttg aataaaaaatc taaattgatc aataaaaaaaa	1140
aaaaaaaaaaa a	1151

<210> 92

<211> 714

<212> DNA

<213> Homo sapiens

<400> 92

ggcacgagta atgcttctgt ttctccact atatgcataat gtatgtgtgg gtacgtgcac	60
atttggtttt ttatttgttt gtgtgtctat ctgaaagttc tgcagggcag cgcttgcct	120
tggattgctg ctgcagtggat gatgagaagg atgagggaaag tgcaggaaaa aggggaggtg	180
ttcaggaaca tggccgcac ctggccctt cgttctggca tacaaagcct gaattctctt	240
gttagctctg cctttttac tattttcatg accttggctt cttcttggaa ctcattgtc	300
tcactttcct cattggtaaa ttggaccggc ctctttctt tctacttctc aagaaactga	360
tgaggattaa tgagatagaa tctggagccc gttttgtgtt aaaaagagtt aaggatgt	420
agaagacgga gtaatgtca tagagaaggg gaacacacat tgcttaccgt gtgatgtgat	480
agagtctcag ggagcacttc tcttcaact gtaactgtt aactagttgg gcaggtggca	540
gcctcatttc tatttgttc tgaagtggat gacatgttag tgcaggatga taggaagtc	600
aaccaaattgc agggactggat ggaatgacga gtcaagattc atggggaaac atctagcctt	660
ctgcattgtc acctgaaaga aacttagcta ttaaaaaaaaaaaaaaa aaaa	714

<210> 93

<211> 810

<212> DNA

<213> Homo sapiens

<400> 93

ggcacgagtc ctgcctcatt tttctgtggg ctcattgg tatgcaagga aataagccaa	60
acagcctgaa agtccagtag aagtgcacgc cagacatccc agtcctccct gttagggttt	120
ctcagaactg ttcccttaag gtctcaggct gcttggaaagg aggctgata gcagaaaaag	180
tggggacact gggactcca aagggaaagac ggcacatggcc tggaaaccggat ttcttcgc	240
ttctggagcc tggctccct atgtggaggg tcatgtggc atggctggca atggtaatt	300
caccgatggc catggagtcc caagttggcc attatttgc ggtaaaagat acattaaccc	360
agatgacctt gccggggggcc agaatagagc ccgtgaggaa ggagagcaag gcaggatcg	420
ccgggaagcg agaggattt tggggaggag caaggcttc cacaggaact ggcacttgg	480
aagtattcac caaggctgt gccatgcgaa accctctta aaggaaccgc atcgtacgcc	540
taacgggcat ttctttta atgtatggt tcagagctat tgtctaccac gcctcgctg	600
cacacgcaca cacacgcaag ttccctcagt cagccgagaa tcctgcccattc tcttttagat	660
aacaaaagct ctttaggcctt atgtttggg taggattgtt cttccatgga caggtattca	720
gttggaaaca agtatatagt cactgcctt atgttatgga gatactccga ttttagtcctt	780
ctgcctctt gggaaaaaaa aaaaaaaaaaaa	810

<210> 94

<211> 1176

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (569)

<223> n equals a,t,g, or c

<400> 94

ggcacgagt	agcttgagga	tgcatactat	aaagccagcc	aatattaatt	aatcttcaa	60
gtgaaagtac	attaccggcc	tccctgcagt	tttaacccta	tacagtgaaa	gtgttagcctt	120
tccttcttcc	aggaattgtt	agcataaaatc	ctgacagttc	cagacagtat	ggaaggatcc	180
cagtagata	ggaaaagatc	cccactcgaa	ggtccaagcc	tagtggata	cctttctgg	240
gcacatgg	ccaagagatg	acttaaatat	ctacaaccac	ttgtcagctc	agtttttttg	300
gggactactc	cagaggtgtt	ccctcacaga	ggcagtggt	aaaaaagtaa	gtagaaaaaa	360
agcagtaaga	cagggatgtt	tggacaaggc	actcattcat	aagaaaggaa	tgatagcaga	420
ttggatgtt	tttggatgtt	ccttatgtat	tgacgttact	gccaatgaat	tttgccttac	480
actgacctt	ttaacgtcaa	aagtgtcaaa	atagatttgt	tgttggca	gttttgtat	540
ggcggttgg	tattattaat	ccgggatgna	ggctggattt	attttttatt	ttaattttt	600
ggcttggctg	acctggaaga	tctactagct	ctctgcctc	acggccaag	gtgtgttctt	660
cccccaactga	cagtgggct	gctgatggct	cccttttaat	tcccatcagc	tgagggctga	720
ctcagtcaac	atcttctccc	catcctggac	cccaagaata	cagaaaaaag	gctcagagac	780
ttagcacatt	attttgttt	taagatgtca	gcacctgatg	tattattta	gtgcttgtt	840
aaatattctg	aaactgtgtt	ttcttttttc	ctttaattta	aatttgtctt	cataaagtgg	900
gcttacaaga	acatttcttt	atcaagttt	tctggatttt	ctgggtcaaa	agtataagtg	960
atttctggac	ttttcttgac	aaaaagtacc	aagaaaagct	gcattaaaac	aacaatcta	1020
attttaaaaaa	cacttagtga	gctaaaacgc	agactcaaac	caaactaatg	aaagctattt	1080
aagagaagtc	agttgaagta	gtttccagaa	tttatttcat	tgtttttca	actctttgtt	1140
aacaccataa	acgtgaatta	aaaaaaaaaa	aaaaaaa			1176

<210> 95

<211> 1028

<212> DNA

<213> Homo sapiens

<400> 95

gcatgatcct	gtggAACACA	gtttggatc	atagatgtga	attaAGACAC	caccgagata	60
cgggctgtga	ggttcatacc	gtgctgatag	cactcgtgg	gtctgtaaa	tgtggtaag	120
acattcaaac	ctgggttga	tactggaaac	tcttcctta	aaactgtgac	catgattca	180
ttcagcccc	ccacacccc	atgtctgcct	tgtttcagag	tgagtttct	atggagccctg	240
tggcccttt	gcagccccacc	tggtggcttc	ttaatgtaac	tcttccttg	gtgcctgga	300
gtggaccact	catctgcagg	cctctcctgc	atggggaggg	taggcaggga	gcagcatgtc	360
tgcaggggt	aaccttgct	cttctgtcag	gctggggccca	ggctgcacca	gccacctgcc	420
acatggtac	agtgccacgg	gccctgcgt	tggccctgc	aaccgtgctc	tggcgggac	480
acctggctgc	tgcaggccaa	ggccgctgtt	cagtgaagag	tcccatgttt	agtatggact	540
aaagtccat	gtttagccay	tgccttcgtc	tccctgtacc	ccagaaacca	ggtcaactgga	600
ccacagtgcc	agatccat	cacgcccgt	agcacctaga	agtgagaaca	ctgtattcct	660
acaatgtaca	cttgatatt	tctccttatt	tagttctag	tgaaacaaat	caagtaagga	720
actatcttta	gttttagatgg	aattatttgt	ttttaattgt	tgccgtattc	atctatata	780
ctaataatttc	aagataagta	atgaacaaaa	cctgtctaaa	ccttttggctt	ccaatgaatg	840
aaagtcatgc	actttat	taggctctat	gtttggctt	ctgcagtaact	tttattatct	900
atacataatt	tggccaaaaa	taagaaattg	gaaagaatga	aatgttttagt	ttatagtaga	960
agaaagatga	tgacactaag	ttgtgaaaat	atgtgtgtat	tttattgaaa	taaactcacg	1020
gcacgtag						1028

<210> 96

<211> 747

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (605)

<223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (642)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (645)
 <223> n equals a,t,g, or c

<400> 96

tcgacccacg cgtccggcca aaggaatcct gagtatgtat gcactttaaa agaaccaga	60
atcatctaaa tattgtcaca tggctcttgc aagtgtatgat aatagtgtatg cttataataa	120
tgaggattag ctgcactcat cagcctgttag aaagtaaaaa gtttctttt cgaaatttcc	180
tttcttgtta agaataaatac ataagtgtta gaaataatag tttcttttaa agactaactt	240
ccttcaagcc ttctctgctc tgtctaata actcttcgtt aagccctatc ctatgtatc	300
gttagatata agggataaag tatattctat gtccgtact ttagccaaga tatttgtct	360
ggacatgttc acaggcacgt tccagctggc agcttatgcc ctttccttgc ttggaaatat	420
tattactttt ctaagtcttt ttgcaagcaa cttttttttt tcctttgttc tctgttgcct	480
ttccctatat agggaaagttt taagttatta gccagtcggg tttaaatttaa attgtgaggt	540
ccagctccag ccaatggaga caggacacaa gctgcataag ggataaaaaac tgcttccctc	600
cttnttcgg gtgtgctgtc accattgttt catctgtgag gngcnccctt tctgcccagaa	660
agtaaaaatg ctttgcgtaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	720
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	747

<210> 97

<211> 628
 <212> DNA
 <213> Homo sapiens

<400> 97

cggcacgagt atatataatat ttttaatata atatataaaa agacttttag atagaattct	60
catattctga tgacctatca cgttgggtgt gcattttaa tcgtgtgtc aaagaaaacag	120
tttatacttag ctctgcagac tatttcaaca tcactcagga gcaaaacaaat tcttatgggt	180
ttaagcagta ctattattgc agattcaacc ttttattttt aatacttaaa aacaggaaaa	240
aggttataag atgtggagct cattggagag taaaattttt ccaacaaaaa ggatatgaca	300
aagtacaarg gaaaacaaaaa ccaaaaaact tcatttatcc caaaaaattttt attttgcgcg	360
taaatgcttt aaaagtgggc aaaaaaggag gtttctcagt agaattttt gcaactaaag	420
gcaaatggaa aactctcaca tagcattaa taagggtttt catgcaatata atccccactat	480
cccaagaaat atctgcagtt caaagctgtc ttttaattttt atgcttccta gtgtttgtcg	540
tttataaaatc ctaaatatttta aagggtgagt tccttaataa tactattcta ataagttacta	600
agacttttct aaaaaaaaaa aaaaaaaaaa	628

<210> 98

<211> 904
 <212> DNA
 <213> Homo sapiens

<400> 98

ggcacgagat cgtcttgcga caagacttgc tgagaagcac cttaaaattt actgtgagcc	60
acattttgtc ttttactgtc tcattcgata gggtagatca atgtccttta ctgttagcaga	120
gactctctca tggcaggac catcatggaa agttctgact acatcaagaa aggcccaat	180
gtctcacctg tgctgggtt caggcagcag gctgtgtatgc cggtgccctt ctgggtggta	240
ctgtgggttct gcttcctgtt atatgtatcc tcacgaagga cttttggatt agccaattac	300
atgccccctac cctgagcttc ttccccagct ctttgacttc ctggacattt gtaaatatcc	360
tgaataagca aaaggataa aattcataga aatatggtgg caaaaatata caacttcagc	420

ccagttctt gggccatgt tggtaaggag tccagttggc aagacaagct gccaaggaa 480
 gtgcctcaga agtctgggtc aaagaggagg gccagatctg ttctgtgaga ccctatgtga 540
 ttgttatatt tttaaataat atataattaa gcaggacaaa ttaaataactc catggcttg 600
 gggaaattgt tgctttaaag tcctggaatg gggctggca cggctggctca tgccattaa 660
 tcccagcact ttggaaagcc aaagtgggtg gatcacctga gtcaggagt tcaagaccag 720
 cctggccaac atgcaaaac cctgtccatg gtgggtgtgcg aggctgaggc aagaaaatcg 780
 cttgaacccg agaggcagag gttgcagtga cctgagattt cgccactgca ctccaaacctg 840
 ggtgacagaa tgagactccg tctcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
 aaaa 904

<210> 99
<211> 576
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (12)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (521)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (535)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (572)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (576)
<223> n equals a,t,g, or c

<400> 99
aaattccccg gntcgaccca cgcgtccggg caacattccg ttacatagag aaatctatgt 60
ataagctgt gtcataaacac ccatcagttg tatttatgt ctttattaat gtattttgtt 120
tttaagatct ttttcagag cctctgtgtc ctgggttact gtataacttcc cttgacagta 180
gcaatgctga tttgccggct ggtacttttgc gctgatccag gacctgtaaa cttcatggtt 240
cggcttttgc tggtgattgt gatgttgcc tggctataag ttggtaagta tgtacttatt 300
tccacaataa cagaacagac aaaaacatga ttaatgatg aagaccagat gaggagcagt 360
ataagtccaa agtagatgt gagtgatatg attcttgata gtattatcca tagaaccctc 420
ttccctgagt aggcaatgt ggggcttatac tgagttggat atctggactt ataagatgtg 480
gagagtacaca tckttttct ttctttaaaa aaaaaaaaaa nggcgccgc tctanaggat 540
ccctcgaggg cccaaagctta cgcgtggcat gngacn 576

<210> 100
<211> 713
<212> DNA
<213> Homo sapiens

<400> 100
 ggaagaggta caagcaaagg ttttacgtat gatcaatgtt tacttttagcg gccctggggt 60
 gctaaccgcct ttggatgacc aaggctcacc ctgcacctccg gcaccctttg ctgccttc 120
 cccttgcctt caccctgctg gctcaggggt gctgtgtgt tgccccctca ggctgtgccg 180
 accttgcagg attttgttca ctgggcact cctgctgacc cttcaccate tgctctgtga 240
 aacctctccg agtgttatag gagttggaaa catagtcctt gggccagac ctttgggtgt 300
 aaatccagtc ttcccttattt cttagctgtga ccttgggcaa gttgctgagc cacttttggt 360
 gaccatttcc tcctgaaaaat gaaactaatg atagtgccta cttcacaggt aaataggagt 420
 atgaaatgaa ttcacatata taaagcttct agagcatctc ttctcagttac ccaacatcct 480
 gattactaat ttgcgggggg tggcactctc tcctctttt ctctgctctt tgcaagggtgt 540
 gccaccacta acaataaaact ataggagga gaaaccagt caattccctg aaaagtctcc 600
 agtgtgacca gaagtacaga taatattgtt ccattgtatt aaagtcatc tagggagttct 660
 tagaagatta gatgcatgtt gttccotaca gaggaaaaaa aaaaaaaaaa aaa 713

<210> 101
<211> 649
<212> DNA
<213> Homo sapiens
<400> 101
 ggcacagggta agtgtcaagg gggcgctccc ccatctccgc cgctattacc actgaacccg 60
 gaccccccac ccaggtccag ggccagccgc catgacgaac gtgtactcct tggatggat 120
 tctgggtttt ggtttgcctt ttgttgcac ctgtgcctac ttcaagaaaag tacctcgct 180
 caaaaacctgg ctgttatcag agaagaaggg tggttgggggt gtgttttaca aagccgctgt 240
 gattggAACC aggtgtcatg ctgctgtggc aattgtttgt gttgtatgg cctttacgt 300
 cctgtttata aaatgaattt caaagcaccc aagtcatcaa ctgccaacca aggggacggg 360
 gatgaagaac ctgttggaga cctgaaccca gtgttaggaga gttcagctga aatcatcggt 420
 ccccaggatg acaccacagc atctgcctt gctatatgtg gggaaaactc atggtcacga 480
 acattatatta tgcttcaggg gactacagaa agccagctt ctttgatcta tttgtaaatc 540
 agtccttggc agagtgcata taatgtccgg ataaattaca cccctcggtg ataagattac 600
 atacctcctt cataaaaaacc tgtaaaaaaaaaaaaaaa aaaaaaaaaa 649

<210> 102
<211> 697
<212> DNA
<213> Homo sapiens
<400> 102
 gattnntagggta gggggctgtg atgtaaaacg tctccctgc caaaggaggg gcaaagtgt 60
 gtgtcagttc ctgtttcttc ccatttcctg gcacactctg cccctctgtc cgggggacac 120
 ggcgcattgtgt ttggccaggaa tggggccacc ggtttatgc caacgcctcg ggtgcctgtc 180
 ttgtctgtgt ggcttctcag atggtgagg gtgtctggag ctggcagggt cttccagac 240
 agtctcagcc tctcccccgc gcccccaaca ggctgtcaaa caaaacccga gaggggggtgg 300
 gggagccagc ctcccagcgt gctgtkcccg caggcacccg tttgtatcc gcacgtccag 360
 ctccgtgacc tttgtgtgtg tttgtgtgc caagttagtg agagatttcg aacgcccacc 420
 cctcgacttt gaaatctgag caaaacaaga aactggggtc ttcccttc cccaaacctct 480
 ccccagctag tctccctct gttcttcctg cttccagcc cccgcgcccag attttggaaat 540
 ctcggagaca aaacttagtac tttttttgt actgtatata ttgtgtataa 600
 cgattnnnnnn aaaggagaat tctgtacatt tagaactctt gtaaattaaa aaccgatcct 660
 tttttaaaaa ctgtaaaaaaaaaaaaaaa actcgag 697

<210> 103
<211> 1288
<212> DNA

<213> Homo sapiens

<220>
 <221> SITE
 <222> (462)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (813)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (825)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (834)
 <223> n equals a,t,g, or c

<400> 103

cgatgaaggg gtacaggag aaagattatt taggatcctg aggataatg gagaaaaacc	60
atacaactt gttgattact ttcaactgtga atacatggc tactaccta accgggcct	120
cagagctact ttcaagcattt tggtttccgt agtttgcttg ctcttcctgg gttccatagt	180
gaactgtttt ttaaatgtat tcttcaagcc actgaccctt aactttcca cgcactctc	240
agcatggaga aaagagtcat cagcctggaa ttcccttggt ctccctaccac caacagatga	300
atatcccaca tgagcatcca ctttcggccc ctttccttagc tcagtggct tccttcctat	360
tagtgtctgg ttttctattt cattcaggcc ctacccttcc tgccctgctca gagtcctcac	420
accttatgtt actgattatc cccttwctg ttttggctt tngttttt gagacaagtt	480
ctcaccctgt ccccaaggta gagtggttgc gcataatcat tgctcaactgc agccttgaac	540
tcctgggctc aagcgatcct cctacactcag tctccctgagt agctggact acaggggtat	600
gccacccctgc ccagcttatt tttcattttt tacagctctt cttttaggtt cagggataca	660
tgtgcaggtt tgttacatca gtaaatgcatt gctgcaggag cttgatgtac agattattt	720
gtcaccagg taataaacat agtacactgtat aggtatttt tgaataccct ccctyctccc	780
accctccacc ttcaagtagg cctcaactgtc tgntrgtccc cttcnntgtg tccntatgt	840
ctcaaagttt agctccact tataagttagt agtatgtggt atttggttt ctgttcctgt	900
gttagttgc ttaagatatg gcctccagat ccaaccatgt tgctgcaagg acatgatctc	960
tgtttttta tggctacata ggattccatt gtctatacgt accacccccc aaaatccagt	1020
ctatcattga tggcattta ggttaatcc atgttttgc tatagtgaat agtgctgcag	1080
tgaacatact catgcacgtt tctttatgtc agaaacatata ttaaactaaa gagttctgc	1140
acagcaaaaat aagctataac agagtaaaca gacaacctac agaatggaa aaaatattt	1200
caaactatgt atccaacaaa gatctaataat ccagacgcta taaggaactt aaacaaattt	1260
acaagcaaaa aaaaaaaaaa gggcgccc	1288

<210> 104

<211> 1027

<212> DNA

<213> Homo sapiens

<400> 104

gtccgccccac gcgtccgtac aatgtatggt gtgtgtttgt gtgtataagg tttgataat	60
ttaactttt ttaaatagat ttatgtatgg tagtaatga tagacttagt tctacatgt	120
ttttatgtac tcttcacata ctttattttt ttttgatatt tctagtcatt gaggttcattc	180
tgtttttca aattgttgca aatctccaaa aaattttcca atacatttt taaaaaaa	240
tccatgtata agtggaccca cacagttca accaagttt gtcaggatt gactattgt	300

ctatctaaac atacctaaac atagraaaagg tacagtaaaa atacagtatt ataatcttat 360
 gggatcacca ttgtctatgc aggctgacat taaaatgtca ttatgtacag catgactgta 420
 tagtgtttcc gagttctgtg aggctctcta gcaaactaat ggagctcaag aaggggttat 480
 gggAACCCta acttatacgct agttgggttag gacccttggt caccatctgg ggcttctgat 540
 tgtcatctga agtgggagcc atcttgcgc actgagcytt caaccatgg tatctgtgc 600
 tatctccggt agtgaatgaa gtgaattgaa ttagaggaca cccagctggt gtctgctgca 660
 aaattgctta tttgcttaat gcgtgggaa cccccctcca cacacatctg gagtcagaaaa 720
 ggggtttgtg agattaaagt gggagaaact gaatttgttt attcctatat tcagaatggg 780
 gtccttgara acatcatagt ggtaaggcata gatgttctaa agtcagactg cctgggttca 840
 tctctctgct ccaccacttc gagagttact ttagctcact gtgcttcagt ttcctattaa 900
 attgggataa taataccatc tcatagagta acttaagaat taaatcagtt aatatacata 960
 aagcacttgg aagtgttga agcattaata aacactcaat agctaaaaaaaaaaaaaaag 1020
 ggcggcc 1027

<210> 105
<211> 710
<212> DNA
<213> Homo sapiens

<400> 105
ggcacgagct gggcctccag gttcttcacc tgtcacatga tcattttaca tattgtggtc 60
tgtttattta ccatcagcat catagaagag caaaaagaag aaatactgtg ctccactaaa 120
agccaggctg agaaaacagt tactcacatt gagcagttag tgaccactag gtgggcattt 180
gttcatagct gcatggagaa caagtggcca tatacatctt tctgctgtatc cagcctctaa 240
attttgaatg catcagttt taaaactgca tttagcaata ttccgtgggt gtgatccata 300
atagcgtaac tatttacgcc tggacagag agaaaaactg tatggatatc agatatctt 360
aagagcttt taatctttaa tcaagtttgtt acttcttaag gatgattaag gccaggcagt 420
ggctcacacc tgtaatcccgcatttggg aggccaagat gggtgatcc cttaaaggcata 480
agagttcaag gccatcctgg ccaacatggt gaaaccccat ctctactaaa aataaaaaaa 540
ttagctgggg tgggtggca ggccctgtt accccagcta ctcaagagggc tgagacaaga 600
gaatcgctt aagccaggag ttggagattt cagttagcc agatcatgcc acttcactcc 660
agcctggaca gcagagtggg acttcttctt aaaaaaaaaaaaaaaa 710

<210> 106
<211> 530
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (16)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (22)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (45)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (47)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (54)

<223> n equals a,t,g, or c

<400> 106

ttggcccccc tagggncctt tngccaaaaa aggctatTTT agggngnccc cctnttagaa	60
gggtaccccc cttcaggtac cggccccgga attccccggg tcgaccccac gsgtccgccc	120
cacgsgtccg cttttgtttg gagaacagct ggctaaggat gactctaagt gtactgttg	180
catttcaat ttggtaaag ratttgaatt taaaatTTT ctttttagct ttgaaaatat	240
tttgggtgat actttcattt tgcacatcat gcacatcatg gtattcaggg gctagagtga	300
ttttttcca gattatctaa agttggatgc ccacactatg aaagaaaatat ttgtttatt	360
tgccctatacg atatgctcaa ggttactggg cttgctacta tttgttaactc cttgaccatg	420
gaattatact tgtttatctt gttgctgcaa tgagaaaataa atgaatgtat gtattttgg	480
gcagacacct gaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa gggcggccgc	530

<210> 107

<211> 392

<212> DNA

<213> Homo sapiens

<400> 107

tcgaccggcgg	cgtccggccca	cgcgtccggaa	gggaacttaa	atgatattcc	ccttttcctt	60
ttccctaata	accccttgc	cagatttca	agtaggcaat	gataaacagca	ggtgagat	120
taggaactgt	gactacgaat	atgttagatgg	agatgtgcag	aaggatccag	agacttagag	180
caatgtttca	catgttttgc	gtaaagcatgc	tccctactgt	ggtaaaacca	aacatgtacc	240
aacccccccca	gaatttatgat	attctactgc	agtaaccagc	ctttttttt	aacatcgat	300
agctaaagga	cgttatcctc	aaagtcatgg	aaaagcagga	agtttttcat	gacaaatcg	360
tttgcctatag	tacagttaaa	aaaaaaaaaa	aa			392

<210> 108

<211> 991

<212> DNA

<213> Homo sapiens

<400> 108

ggcacgagga	attttgtcac	gtgagctgtt	gggttactga	gtgagtgaaag	ttcactgtct	60
gcaattgagc	ccttttgagg	attctaaaaa	cttcagcctt	ttcagtcctt	tccatctcat	120
tcccccttaaa	gaaacacatt	tggactttgt	ctggctctct	ggtaaaaccct	gtgacctgca	180
ttacttagggc	acagtgacac	agaaaaagaaa	gtgtgtgttt	ggtaaaatatt	tattgagcac	240
ctgcagtcctt	atgttttagca	gtatgcatgg	tgccctgcctt	tggaaagcaa	agcaaaccag	300
ttcatcagca	ggttttcttt	gcctgcatgt	sctgtgccta	gccttgcagt	tgacacgaga	360
gaaatataaa	acatggccct	ggccttcctt	catttaaaca	tttctytttc	ccaagcactt	420
actctgtgca	aggagctgga	gaagccaaag	ctggagaaaa	acaaaggagg	gcctggccctc	480
gagaaggtag	tgttctaatg	ttgtggttct	caaactcagg	cgtgcgtttt	aatcgctctgg	540
gggccttgcc	agaccacaga	acccatcccc	tgagtttctg	aatcaatggg	tctgaggttg	600
ggctcgctga	gaatttgcat	ctttataaaat	tccagataat	ggtcttgcgg	ctgggttaggc	660
accatggttt	aagaaccact	ggtctggccg	ggcgcggtgg	ctcacgcctg	taatcccagc	720
acttcggag	gccgagacgg	gcggatcacg	aggtcaggag	atcgagacca	tcctggctaa	780
cacggtgaaa	ccccgtctct	actaaaaata	aaaaaattag	ccgggcgtgg	tggcgcgcgc	840
ctgttagccc	agctacacgg	gaggctgagg	caggagaatg	gcatgaaccc	gggaggcgg	900
gtttgcagtg	agtgcagatc	gcgccactgc	actccagct	gggcgcacaga	gcgaaatccg	960
tctcaaaaaaa	aaaaaaaaaa	aaaactcgtat	g			991

<210> 109
<211> 912
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (896)
<223> n equals a,t,g, or c

<400> 109

gggtcgaccc acgcgtccgc ctcaagggtgc ctactttgct ggttcccttt ccagcagctc	60
ccccacacctcc cttagcccccc cctcctctgg cagcctctcc tgcctctgct gagctccct	120
ccacgtgttc cacccttta ccctgctgtt gttacatcc aacctgcctg agaatttctt	180
ctggggagga atctattcct gtcatggct agtgcctggg agggagagaa ctttctgggg	240
gtagggtgcc ctccatctga aacaggccag gtgagcatca tgcaayaaggc ctccattctg	300
tccgctcaga ttctgggtgg ggcccacaggc aaatctcctg acttatgggg agttggcttg	360
tggttccccc cttggatagc ctccatggaa ccactatagg ctttcccaac agctgcctct	420
gaaatagctg ctgccttcgag atcctccctt tttaaagcac tttctaaagc cctcaggatg	480
gcgggagcra acagcactgg tatattcttag gagtaagtgc aggaattcag cagttagagc	540
atgtctggga ccacctggac tgccatccat ttaacacctaa atctctttgg gataactcgcc	600
ctccctggga accagagttc tggctctaacc attgagcagc tatgcactag ttccagagaaa	660
gccactaaca ggctgccatg tgttagatgt ggttcttaag agatcacagg ctgggtcatc	720
tgatcaactgg atggatagct cagcctgggg catttagtgt tttccctgg gataaatccc	780
caagrcagct ggatttggag ctggtggcaa gttgaaatta taaaaatttgg atttgtgtgg	840
gactgtcaaa aaaaaaaaaa aaaaaaaaaa aaaaaactcg aggggggggc cggtanccaa	900
ttcgccctat ag	912

<210> 110
<211> 875
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (66)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (872)
<223> n equals a,t,g, or c

<400> 110

ggcacagcgc gaggctgggt cccggcccaag gagaaggaag tcgctgaagg cagtgccat	60
gctggncgtg gaaatggag gcggttgcag rgggtctatg gggcccggtc ctggatactc	120
ggcaggaagc cgtgtctgca gaggctctc cctgcctcag gtggcccccgt tcaaccccaag	180
ccgtgcccattt ctcctgccac cgcctgtcgg tgggggttta aattcggtgt ggcttctgg	240
ggtgcaagctc agcaccctt cttatgcaga ctggggagggg gtcgggcagt cccctcagcc	300
acgaggaccc tggatgggtt ctatgtact tgggaccgtg gggcctggct gcgtactgag	360
tgggtgcccc acagtcaagg ccaacggggg ctccccctgc tctgagatgt tgggagaaaag	420
gcggcttctg gaaccttccg tgggaccgtt aagtggctgt ccagaaaggc gggagggtgg	480
gcacggggca cggggggcag ctgggggtcgt cgttaagggt cacgcattccg tacagttgaa	540
tttcctttctt cttatcatgt ttatccacc ttgtcccttt tttcccaat tgcgttttg	600
catttttttc cttggcaaat gtaaaactcg ccttcattc atgacgtgtg aaatttcagt	660
ttctctggag tttgtcagac ggcgtggaa ccacgcctga aactcaggttta ataggaggaa	720

aaaaaaaaaaa cttaaaaaaaaa acataaaaact actctctacc gctggsc 780
 cagcctgtct cgccctggcc gcggcagggt ggccctgtaac aatttcagtt ttcgcagaac 840
 attcaggtat taaaaggaaa aaaaaaaaaa anggg 875

<210> 111
<211> 459
<212> DNA
<213> Homo sapiens

<400> 111
 gggtagagaga gggaggtacc agagtaaata cagtgccact tggatggta cacccatta 60
 ctttcagaca caaagatgta agctgaggaa aatgaattct tggattcagg gaaatgaatt 120
 cttggattct gaacatgagg gtcagattt cattcctgtc tcaattgttg acgcttatcc 180
 caaggactag tcattctgca acatctgtgg gaaattccca gattgaactt cccagggaga 240
 aacatcatat gacatattgg gaaaatggct gacaatggtt ttccttagta agttcattga 300
 gaagaaaaagt gggcggtatga tttcccagt cttcactctt tcagaagccc ctaaaacaat 360
 ctgacatgct ctatgtggag cctgctttct atcccatcag tttgatttct gaatgcctta 420
 tgatcattag cattttcat taaaaaaaaa aaaaaaaaaa 459

<210> 112
<211> 609
<212> DNA
<213> Homo sapiens

<400> 112
 ttctgttcat ttttccccct ttctcccca cattcattaa gaaccctact gaaaccctag 60
 gtgacaaaag gtgtgccttc tgttgccaca tttgaccac cacaggactc actggactgg 120
 acttctattt atatggatt aqtaactga tatatatata tatataakt tttgattgac 180
 accaaaaaat taccttggca caaatgccag acctgtgaag gtcagaggcc cgctgcttyt 240
 cccaggaggg agggacttt ttgggtgtct gtggcaattc ctctgtacag attgttaactt 300
 tttaaaaatt tcccttcacc ccgtcaactt aatatatgtt catagtaatt tgtaagatac 360
 ttcttttctt tattttgggtt gcaagaccct tccgaacaca ttcctgtata aagtattttg 420
 cactattnaa agaaaccat atggatgaag tcaggatgtg caatatgtt gcgtcacagt 480
 gctcatcggtt gtacctgtaa tgtaactaat cagtttaat gtactattt aatatgtaa 540
 aataaatttt caccatgagc atgttttaat gaaaaaaaaa aaaaaaaaaa aaaaactcsa 600
 gggggggcc 609

<210> 113
<211> 1404
<212> DNA
<213> Homo sapiens

<400> 113
 tacgagtttt tttttttttt tttttttaaa ggagagggtg caatgtgtt ccctagctag 60
 taagagtccc atctggcct tctaaggaa agataggtaa atgaaaagac tgctaaatcc 120
 aaggtcagac agcatataga aggctttata aaagaacagg aaaactcaga acactaaata 180
 agagagtgtt ttcttgact ctgcagttgg cctcaatcat cggatctggaa atattacttt 240
 ttacgatttt gsaagccgat acacacctgt aagtaataac tgaggaaggt agagtatgtat 300
 tagtctttt accttcagt gtgtatcaat gttaagtgaa caagagcmma aggaaaaacca 360
 tatattttgtt attttgcaac atatataaaa taacaacact gggctggcg tgggtggyc 420
 agcctgtawt cccagcaytt kgggagccga ggcagggggta cacaagkta aggagttga 480
 gaccagcctg sccaacatag taaaaccccg tttccactca aaatacmmaa aatttagtcgg 540
 gcgtgggtat gggcacctgt aatcccagct actcgggagg ctgaggcatg atgatcgctt 600
 gaacctggga ggcagagggtt gtcgtgagcc gagatttgc cactgcacac cagccggga 660
 aacagtgcga gactccgtct taacatgaaa aacatgaaca gccgctacta tctgaggggca 720

atttttgtc ttat	ttt ggcattgtata ttat	ttcac aaat	ttt aaaggccagg	780
tgcggtgct cacgcctgta atctcagcac	tttggggaggc cgagggtggc ggatcacgag			840
gtcaggagat cgagaccatc ctggccaaca cagtgaacc ctgtctccac	taaaaaaacac			900
aaaaaaattas ccaggtgtt ggtggccggc acctgcagtc ccagctastc	aggagtctga			960
ggcaggagaa tggcatgaac ccgggaggcg gagcttgcag tgagccaaga	tggcgccact			1020
gagctccagc ctggcaaca gagcgagact atgtctcaa atagtaataa	taataat	ttt		1080
aaaataaggg gaaaaaaatc actgataaac caaaaaccc aacctaaga aacgttcaca				1140
tctgtatagc taatactctg acgatggga tacaaaaaca ctttcaactca	gtggcttgc			1200
agatatcatt ttttcccag tatttttgg aaagaaccaa tctttgtctt	ttttctcct			1260
tcttcaggg actttatgaa tccagaaaga gccaacgtt gaatgattac tgcaatctca				1320
catctattaa atccgtatac ctgcaaccaa gagatgagta ggagatgtgg atcctaagag				1380
gtgacctgta acatactgcc ccct				1404

<210> 114

<211> 853

<212> DNA

<213> Homo sapiens

<400> 114

gggtcgaccc acgcgtccgg gtgaattaac acgtacccaa tggccaagag tagat	tttggg	60	
tgtcagtat aaaaat	tttcaaaaaa cctgggtttc tcagttacag ct	tttatataa	120
gtatagtaat aacttagca gagctgtaga gagatagatt tgcaaaactt	g aagtgtat	180	
ggataaaatct ccatacgtgg tagaatttata tataaaatgg catattcaaa	ggtatgtgt	240	
attat	tttgcatttctgaa gaaacttagta tcataaaaaa	tgttgcgtat	300
ctgacatcag aattccagaa ttcatatgcc accccctgtt	ctgggcttcc	360	
gtggcttggaa ggggtggtgc tgggtacggg tgggtgaggc acgcatgca	ggtattgcag	420	
aaggaaccca cgcaaccgtc atcccttctt ccccccaagtg atgctgcctc	attctgggt	480	
cctgaaagta ggcttca	ttt aacatgttag ggaagtttctt ggctgaaaaa	gcaaaaggct	540
tttattcactg gagtctatcc tgagccccct	gtgcaaaagg cagtgtgaac tcaggggaca	600	
gaatcactga agctttgtaa aagcacaac atctgcctat	cacagtccaa aggggacttc	660	
aaaatcaaga atgtctgtga cggagaagat gggaaacagaa	cctggctgtat	720	
gaatcttctc tgggtcgaga tggatcgtt gaccgttttcc	tttatttcat	780	
ttttaatata ttccacccccc tgcataatatt ctgtttactg tggat	tttattgtt	840	
aaaaagggcg gcc		853	

<210> 115

<211> 845

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (845)

<223> n equals a,t,g, or c

<400> 115

aactagtggaa tccccccggc tgcaggaatt cggcacgaat ggatctgtgt	ggatggtgt	60
tgccttgggt gtgtatgtmt gtrtgcgtgc acccaactgca	gcagtacccaa	120
gggcaccatt tgcttgaaga tgcttctgg tgccaaactgt	gcctgccaag	180
cagacagcat tagacaggct gtgacgttgc caggcacggc	cagagccaag	240
cagctccttc tccagctgtc actgctccca	gccttctgt	300
cccaaggcct tcaggctgac ccctggattt cagaacaccc	ctttcatca	360
aggctgggtt ccattgtccat caacagcaag acaccgttc	ccctgcataa	420
aagtcaggag gactcaggca gacacactgt	acacatcacc	480
cccaacgact wtaagtgggg tggatatag cctgtttac	gcaagcttc	540
aagaagttaa gtgat	tttgcagcag ttttgggtt	600

gatggagaga agttacacca ggacttagtg cctgggaagc aaagaggtga aactactcg	660
ccaggattgc acagctgaca gtgatgatac cgatggctgt gcttttagta gctgttaggt	720
accaggaact gtgcggcc cttgacacat ataatttac ggaatcctca cagcagatta	780
aaaaaaaaaa aaggtaaccat attgtccccca ttttacagac cacccttac aagagtggga	840
tggtn	845

<210> 116
<211> 760
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (13)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (300)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (425)
<223> n equals a,t,g, or c

<400> 116	
cggacgcgtg ggnccggacgc gtggggaaaa aataacaaaa caaaaaacaa gaaaaaaaaa	60
acacaaaaacc ccgtaaaatc acaaagaaaaa tccaacacca aaggcgcaga agccggctgg	120
ccgtgggtgg ggcagcgtag gcgtasatcc ctctcccttc acttagcctg ttgactcttg	180
ttattatcat gatattcaca aaacgccgca tggtaaaaaa gtcataagatg tcatcttctc	240
tctgccccca gggaggaaag ccaccttctc ttgcccccttg gcccccttgc cagggggccan	300
gggtctgcgc ggtgggggtg ccaacaggcc tggcccttgc ctcccctgca tccagccatg	360
ggggcctctg cgattgcccgg aaggttgcatt ggcgtgtccc agggcagca caggcccgag	420
gccngngctgc ctggtttat ttttattttaa ctttattttc tgtttatgatg gtgtgtgtcc	480
gcccacccccc acccccttca gtgttaagtgc gggagccctg gggagatctc tcctgcctcc	540
cagccctctcc caagacctcc cccctcgatca ccagccatcc ctctggacca ggcagaggc	600
gacccgggtg ggcaggggcc tgggggtggc tggggccagc ccaccagcca atggaccct	660
cctcaggccg ccagtgtcgc cctgccccctt tttaaaacaa aatgcctctg tttgtaaacc	720
cttagacgct tgagaataaa ccccttccctt ttcttccaaa	760

<210> 117
<211> 988
<212> DNA
<213> Homo sapiens

<400> 117	
gtacgacgtt ggctccctg gctcctctct gcacccctccc agcaatatgc	60
atcttgcacg tctggcggc tcctgctccc tccttctgct actggggggcc ctgtctggat	120
gggcggccag cgatgacccc attgagaagg tcattgaagg gatcaaccga gggctgagca	180
atgcagagag agaggtgggc aaggccctgg atggcatcaa cagtgaatc acgcacggcg	240
gaaggaaagt ggagaagggtt ttcaacggac tttagcaacat gggagccac accggcaagg	300
atgtggacaa aggcttccag gggctcaacc acggcatgga caagggtgcc catgagatca	360
accatggat tggacaagca gaaaggaaag cagagaagct tggccatggg gtcaacaacg	420
ctgctggaca ggcgggaaag gaagcagaca aagcggtcca agggtccac actggggtcc	480
accaggctgg gaaggaaagca gagaaacttg gccaagggggt caaccatgct gctgaccagg	540

ctggaaagga aktggagaag cttggcccaa gtgcccacca tgctgcggc caggccggga	600
aggagctgca gaatgctcat aatggggtca accaagccag caaggaggcc aaccagctgc	660
tgaatggcaa ccatcaaagc gnatctcca gccatcaagg aggggccaca accacgcccgt	720
tagcctctgg gcctcggtc aacacgcctt tcatcaacct tcccgccctg tggaggagcg	780
tcgccaacat catgcctaa actggcatcc gccttgctg ggagaataat gtcgcccgttgc	840
tcacatcagc tgacatgacc tggaggggtt ggggtgggg gacaggtttc taaaatccct	900
gaagggggtt gtactggat ttgtgataca ctaaaaaaaaaaaaaaa	960
aaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa	988

<210> 118

<211> 1947

<212> DNA

<213> Homo sapiens

<400> 118

gaattcggca cgagttgtgg tctatttaat gccatgcttt tcctgtttt gtactgtttg	60
ttggttgttt tgccatttaa attaaccccc aagcatagtg ctgaagtgc gcttagcatt	120
cacaagtcca agaaatattt atgtaaagtg aaagctgcct gcaagattca agcctggtat	180
agatgttggc gggcacacaa agaatatcta gctatattaa aagctgtttaa aattattcaa	240
gttgccttct ataccaaact agagagaaca cggttttga atgtgagagc atcagcaatt	300
atcattcaga gaaaatggag agctatactt cctgcaaaga tagctcatga acacttctta	360
atgataaaaaa gacatcgagc tgcttgggg attcaagcac attatagagg atataaagga	420
aggcaggctc ttctcggca gaaatctgct gctttgatca tacaaaaata tatacgagcc	480
agggaggcyc gaaagcmtga aaggataaaa tatattgaat taaaaatct acagttatcc	540
tacaagcayt ggtgcgttgt tggtagtac gaaaaagatt tttagaacag agagccaaaa	600
ttscgacttc cttccacttc actgcagctg catattatca cctgaatgct gttagaattc	660
aaagagccca taaactttac ctggctgtga agaatgctaa caagcagggtt aattcagtca	720
tctgtattca gagatggttt cgagcaagat tacaagaaaa gagatttatt cagaaatatc	780
atagcatcaa aaagatttag catgaaggc aagaatgtct gagccagcga aatagggctg	840
catcagtaat acagaaagca gtgcgcatt ttctctccg taaaaagcag gaaaaattca	900
ctagtggat cattaaaatt caggcattat ggagaggcta ttcttgagg aagaaaaatg	960
attgtacaaa aattaaagct atacgactaa gtctcaagt tgttaatagg gagattcgag	1020
aagaaaacaa actctacaaa agaactgcac ttgcacttca ttacctttt acatataaagc	1080
acctttctgc catttttag gccttaaaac accttagaggt agttactaga ttgtctccac	1140
tttggttgtga gaacatggcc cagagttggag caatttctaa aatattkgtt ttgatccgaa	1200
tttggtaatcg cagtattctt tgtatgaaag tcatcagata tgctgtcaa gtctgttca	1260
atgtatctaa gtatgagaaa actacttcag cagtttatga tgttagaaaat tgtatagata	1320
tactatttggc gctttgcag atataccgag aaaagcctgg taataaagtt gcagacaaag	1380
gccccggcat ttttacaaaaa acttgggttt ttgtggctat ttactgaag acaacaaaata	1440
gaggcctctga tgtacgaagt aggtccaaag ttgttgaccg tatttacagt ctctacaaac	1500
ttacagctca taaacataaa atgaataactg aargaataact ttacaagcaa aagaagaatt	1560
tttctataag cattcctttt atcccagaaaa cacctgtaa gaccagaata gtttcaagac	1620
ttaagccaga ttgggtttt agaagagata acatggaaaga aatcacaaat cccctgcaag	1680
ctattcaat ggtgatggat acgcttggca ttcccttatta gtaaatgtwa acattttcag	1740
tatgtatagt gtwaagaaaat attaaagcca atcatgagta cgtaaagtga ttgtgtct	1800
ctgtgtwcaa cttttaaaat ctgactttgt tttaaaaaaaa cataaactgt tcattacatt	1860
cttcattttt atcatttata gtttatgca tgtaataaac taatatgtca taagatgaaa	1920
aaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aactcga	1947

<210> 119

<211> 1448

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1441)

<223> n equals a,t,g, or c

<400> 119

tcgacccacg	cgtccggAAC	gtgtcccgCG	ggctcAGTCC	gcccGCCGt	gcgtCCGCG	60
agtcaagtg	agtttctCGG	ctgccccCG	ggccGGGGt	cggagCCGAC	atgcGCCGc	120
ttctcgGCt	cTTCTGGTC	ttcGCCGGt	gcacCTTCG	cttGTACTTG	ctgtcGACGc	180
gactGCCCG	cggcGGgA	ctggGCtcca	ccgaggAGGc	tggaggCAGG	tcgtctGTTG	240
tcccCTCCG	cTTGGCAGAG	ctgcGGGAGC	tctCTGAGGT	cTTGAGAG	taccGGAAAG	300
agcaccAGGc	ctacGTgtTC	ctgtCTTCT	gccccGccta	cctctaACAA	acaaggCTT	360
gccatCCCCG	gttccAGTT	cctGAATGTT	ttAGCTGGT	cTTGTTG	gccatGGCTG	420
gggCTTCTGc	tGTGCTGTGT	gttGACCTG	gtggGTGCCA	catGCTGCTA	cctGCTCTCC	480
agtattttG	gcaaACAGTT	ggtGGTGTCC	tactTTCTG	ataaaAGTGGC	cctGCTGCAg	540
agaaAGGTGG	aggagaACAG	aaACAGTTG	tttttttCT	tatttttt	gagactTT	600
cccatGACAC	caaACTGGT	cttGAACCTC	tcggccccAA	ttctGAACAT	tcccatGCTG	660
cAGTTCTTCT	tCTAGTTCT	tATCGTTTG	atcccATAATA	atTTCATCTG	tGTGCAgACA	720
gggtCCATCC	tGTCAACCCt	aACCTCTG	gatGCTCTT	tCTCCTGGGA	cactGCTTT	780
aAGCTGTTGG	ccATTGCCAT	ggtGGCATTa	attcCTGGAA	ccCTCATTAA	aaaatttagt	840
cagaaACATC	tgcaATTGAA	tgAAAACAAGT	actGCTAATC	atataCACAG	tagAAAAGAC	900
acatGATCTG	gATTTCTGT	ttGCCACATC	cCTGGGACTC	agttGCTTAT	ttGTGTAATG	960
gatGTTGGTCC	tCTAAAGCCC	ctcATTGTT	ttGATTGCC	tCTATAGGTG	atGTTGGACAC	1020
tGTGCACTAA	tGTGCAgTGT	ctttTCAGAA	aggACACTCt	gCTCTGAAg	gtGTATTACA	1080
tcAGGTTTTC	aaACCAGCCC	tggTGTAGCA	gACACTGCAA	caGATGCTC	ctAGAAAATG	1140
ctGTTTGTGG	ccGGGCGCGG	tggCTCACGC	ctgtAAATCCC	agCACTTTGG	gaggCCGAGG	1200
ccGGTgATTC	acaAGTCAGG	agtTCAGAC	cAGCCTGGC	aAGATGGTGA	aatCCTGTCT	1260
ctaATAAAA	tacAAAAATT	agCCAGGCgt	ggtGGCAGGC	acCTGTAATC	ccAGCTACTC	1320
gggAGGCTGA	ggcAGGAGAA	ttGCTTGAAC	caAGGTGGCA	gaggTTGCAg	taAGCCAAGA	1380
tcacACCAct	gcaCTCCAGC	ctGGGTGATA	gagtGAGACA	ctgtCTTGAC	aaaaaaaaaaaa	1440
aaaaaaaaaa						1448

<210> 120

<211> 496

<212> DNA

<213> Homo sapiens

<400> 120

tcgacccacg	cgtccGAAct	gacacaATGA	aactGTCAGG	catGTTCTG	ctcCTCTC	60
tggCTCTTT	ctgCTTTTA	acaggGTGT	tcaGTCAGG	aggACAGGTT	gactGTGGT	120
agttCCAGGA	cACCAAGGTC	tactGCACTC	ggAAATCTAA	cccACACTGT	ggCTCTGATG	180
gccAGACATA	tggCAATAAA	tGTGCTTCT	gtAAAGGCCAT	agtGAAAAGT	ggtGAAAGA	240
ttagCCTAAA	gcatCCTGGA	aaATGCTGAG	ttaaAGCCAA	tGTTTCTTG	tgactTGCCA	300
gtCTTTGCAg	cCTTCTTTCT	tcactTCTG	ttataACTTT	gCTGGTGGAT	tcTTTAATT	360
cataAAAGACA	tacCTACTCT	gcCTGGTCT	tgAGGAGTT	aatGTATGTC	tATTTCTCTT	420
gattCACTG	tcaATAAAAGT	acATTCTGCA	aaAGCAAAAA	aaaaaaaaaa	aaaaaaaaaaaa	480
aaaaaaaaaa	aaaaaa					496

<210> 121

<211> 1174

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1151)

<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1160)
<223> n equals a,t,g, or c

<400> 121

gagagggttca	ggttggtctt	gctggcatct	tcacctaaaa	tagggctgg	cagacaggcc	60
catgcgctgc	atcctgccc	cggcctgata	gatggcctgt	aatggctt	tttacattt	120
aaaagtgggt	aaaaacaaat	taaaaatata	tttcatgaca	tgaaaatcat	gaaattcaca	180
tttcattatc	tgagaataaa	cttttatgga	tgagccatg	cctgtctgtc	catgtcttat	240
ctgtgtctgc	tttgtctac	ggctgcagag	tggagtggct	gggactgaga	cggAACGCC	300
tcctcacggg	gccgcgtctc	tccaccagga	ccgcggggct	actctgaggc	tctgcttctt	360
ccccagcggg	gttggcttcc	tgcttattcct	cagtatctg	ccttggctct	gagttggtcc	420
ctctgscaag	agccgtttct	gtgtctcagt	ggatggcgca	ctgsccttct	tgttggtacc	480
ttgactgata	gackggttcc	tgttcackgc	yccgaagtca	tcccagaaaa	cctctyacag	540
ttgcatgggt	tgaaccagt	ccgcgtgtat	ttagagttt	gtctcttgcc	ccttcaccca	600
gaacagcagc	acccaccacc	ttccctgtccc	ctgtgactgc	ctcgcaactg	ggtctgttct	660
gtgagatgtc	gccaccctgt	ttgccatctg	ggaggatctc	actcctcaa	ttaatctgc	720
tctcttcgg	tatttttta	gtttctatgt	attttacttt	taggacattc	cagcctgggt	780
gacagagtga	cggctctaaa	aaaaaaaaaa	aaaaaaaaaaag	cacaccagt	tcttccat	840
ctcttttaat	cataatcatg	ctttaaaaaa	taccctcgag	catatggagc	aaatttaaga	900
taattgttcc	ttttctgcta	attcattatt	actgtcatat	ctaggtctgt	ttctgtcgac	960
tgtggaccac	ttatgtgcga	tccgtggacc	acttgcgtgc	gatctgtcgg	ccgacgatga	1020
gcttgttcgg	atgtagctcc	atcgtaagtc	gaggagcatc	tgtgatttgc	cctctgctta	1080
tgggatatgt	ttttcccgcta	ctragtctgt	gtagtaaatt	tttgactagg	aaaaaaaaaa	1140
aaaaaaaaact	nggggggggn	ccccgttaacc	catt			1174

<210> 122

<211> 1046

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (14)

<223> n equals a,t,g, or c

<400> 122

tgcaggaat	tcgnacacgag	cactagtagc	tgggtgttca	ggctggcggc	gctcacctt	60
ctcctagccg	ggtagcccg	gggattttatt	ttatgttgc	tttctctgaa	atgccaaagc	120
cacccgat	ttcagagctg	agtacttctt	taacgcttgc	cgtggaaaca	ggaagattt	180
cgggaccatt	gcacagagca	tggagaatga	tgaacttccg	tcagcggatg	ggatggatgt	240
gagtgggatt	gtatctgtt	gccagtgcag	cagcattttt	ctatgtttt	gaaatcagt	300
agacttacaa	caggctggcc	ttggaacaca	ttcaacagca	ccctgaggag	ccccttgaag	360
gaaccacatg	gacacactcc	ttgaaagctc	aattactctc	cttgcctttt	tgggtgtgga	420
cagttatttt	tctggtaacct	tacttacaga	tgtttttgtt	cctatactct	tgtacaagag	480
ctgatccaa	aacagtgggc	tactgtatca	tccttatatg	cttggcagtt	atttgcata	540
gccaccaggc	atttgtcaag	gtttctaattc	agatcagcag	actacaactg	attgacacagt	600
aaaatcagtc	accgtttttt	ccctacgatt	acaaaactgc	cagtctata	tggagtctga	660
tcacaagact	gcagtttctt	cacagatctc	aggaagttgt	cgtggggcag	aggcttttta	720
aaaacatgtg	attagggagc	tatctttatc	tgaataataa	cgaatttttta	gtaaaaaccc	780
gagatagagt	actacaaaat	catgttgatg	acttcagatt	ttggaagttt	aatcatgtct	840
gttatttgca	ttcttttagaa	acttgactaa	gtacctgaat	tcatatttct	attctactgt	900
gcaacatagt	gatgattcag	aaattttcc	tttggggaaa	aaaatgaata	tgaacatttc	960
cattgtgtt	agtgtaaaaa	ggtccagaca	tgatcataaa	atttaaattt	tatacaaawa	1020
aaaaaaaaaa	aaaaaaaaac	tcgttag				1046

<210> 123
<211> 1160
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (325)
<223> n equals a,t,g, or c

<400> 123

ggtcctatgt	gtctataact	tatcagattg	ggagcttagca	gaaagagata	agattattgc	60
tatataattt	ttagggatag	acaatttaat	ttcccttggtt	tcctgtctca	tggacctcac	120
ttaaccagta	gtatgtgggt	gtttttctca	ccttttttct	ccatcttatt	taaaatttgt	180
tggtgtattt	cccttagtca	aactaaagag	aaaacaatcat	ctaatcttat	gttttatttc	240
ttttgtatat	gtacatatga	gaggaggagg	aagaaaagaga	tgaggagagg	tgaaaagaaaa	300
agatccttct	gcctgattgg	gcttngccag	catatgatag	cagtgcaggc	ctgggttccat	360
gagcagcatc	agattcaaat	ttcatagaaa	aagagcccag	aggaattgaa	aaagagaaaat	420
taaattcaac	aaggagaggc	attgtataca	ttatgcattc	acgataggtt	atgattgaga	480
agaagctggt	gctttgggaa	aaacatatta	ggttctacat	ttacctttt	tgaatagttt	540
tctcctttct	aaacagggtg	ataataggag	aatgctgaat	gcctctccat	tgaatttgga	600
aactgccggg	ccagcattag	tgtggtattt	tctgcccaca	cttttctaga	tgcaagttta	660
agatcatgtt	cagtgtgaac	attgaggact	ttagagatcg	gagtccgaaa	tgtgtcaaag	720
ttaatgttaa	tagatgtgt	cctcatttt	taactgtgac	ttctaaatgt	gaccttttag	780
ttcatatctc	ataaaatttgc	catttaagaa	gaaatacaga	watgaaaagt	tnaagtttta	840
ataaaaagtat	atcttgctgg	gtgcagtgcc	tcatgcctat	aatcccagca	ctttgggagg	900
ccaaggccgg	cagatcactt	gaggtcagga	rttgaract	agcctggcca	acatggcaaa	960
accttgcctc	tactaaagat	acaaaaaaaaa	tagtgggca	tggtggtgca	catccgtatt	1020
cctagctact	tggaaaggctg	aggcacaaga	atcgcttcaa	cccggagac	agaggttgca	1080
gtgagccgtg	attgcaccag	tgcactccaa	cctgggtgac	acagtgagac	tgattcaaaa	1140
aaaaaaaaaaa	aaaactcgag					1160

<210> 124
<211> 893
<212> DNA
<213> Homo sapiens

<400> 124

ggcacgagta	aggataaaag	tgggcctgag	cccagtacat	cctctgcagg	aggctgaagt	60
ttctgaaaaca	agaagtggga	gagggttcag	taggaagtc	cacaagttag	gtcgaccaaa	120
gagatcctgc	tgtttccccca	tgagtgccac	aaggactgg	ggtggaaagg	ctgaggctgg	180
accagtcctg	gatcagtggt	cctttctgt	gtgttcttc	tctgtcccc	caggtctgg	240
gggctgggag	cctgctgcgt	gctctggaaac	tttactcagt	cttgttgagc	cactttctt	300
gggaaatgtg	gaccatgtct	cttaaagaac	cagaattgt	tctttccacc	aagtattaa	360
ctgtgtggag	arggagagag	cccctgtcag	aaattggggg	atgcagactg	aacaatgaag	420
gaacatagca	acaatgaagg	aacataggga	caatgacwcc	accttgagtc	cagtggatg	480
aggtgcggct	gcattaaaga	atgaggaamg	ggacagagac	aggtgttaga	gacgatggaa	540
caatcascca	agaaagtctag	ggggttggct	ggcgccgggt	gctcacacct	gtaatcgccg	600
cactttggga	ggctgtggcg	ggcagatggc	ttaagcccag	gagttcgaga	ccagcctgaa	660
caacatggca	aaaccccatc	tctacgaaaa	atgcaaaaat	tagccaggca	tggtggcatg	720
cccatgcagt	ctcagctact	tggaaagctg	aggtgggagg	atggcttgag	cctggcaggc	780
agaagttgca	gtgagacaag	atttttaaa	aggccaggca	tggtggctaa	tgccctgtat	840
ctcagcactt	tgggaggcca	aggtaggcgg	atcacctgag	gtcaggagct	cga	893

<210> 125

<211> 1049
<212> DNA
<213> Homo sapiens

<400> 125

ggcacaggaa aaagccatct aaggtgctaa	gttaaaaagaa	aaaaaaaaagg	cottataagg	60
tactcaggat ctacacgagg ttgttaattc	atgtttgct	ttaatttgtt	actctgtttc	120
ctatccctc ggtttccaat acttgggt	agaaaacatc	agttttgtgt	gtatttgctc	180
ttggcctcaa agttaaggac actatacgca	gagttattg	accttcattt	gtgtgccagc	240
attctgggt gacataatgc ctgcaagtgt	catattctta	atatgtgagg	gggttctata	300
tggagtagacag ggtagttgc taaataacct	ctgtaccctt	ttttctctgt	ctcgatgtat	360
gcatctcattc tcctgttagat tgccttattt	tgatgtatcc	tagaaaaggc	cttcgatagg	420
acgtctgttag ggktattccc ttctaaagg	aatggttata	ccctctgacc	tatcaattcc	480
atttctataa atttatccca tagatatact	cacaatgtg	taaaaatgaag	tatatttgaa	540
gttaaattatt aaagcmtnaa gagcaagcta	aatgttcatc	agcagggaaat	ggagtcataa	600
tatcttcgtc tgcgtgtata atgaaacaat	atgtattatt	atgaacagtt	tttgagcaaa	660
taaaaataaag ctgaagttta aaaagttgag	ttaaaaaaaggc	aaggtgtaaa	acagtatgcg	720
tagtatctgt gtacggttgt agataactgta	cacacatgtt	agagggcaat	ttggataaaag	780
tattctgtgc tcaattaaca tattttccct	ttgtcttccct	ggctctactg	gcattattacc	840
agtagcagtt actcggggagt tacccagcta	ctcaggaggc	tgaggcagga	gaatcgcttgc	900
aaccggagg gtggagggtt cagtgagctg	agatcgcc	tgggtgacag	aacaagactc	960
cgtctcaaga aaaaaaaaaatg cttatgtct	gtataaaaatc	ttcaawaaaa	tgacgataacc	1020
agtaaaaaaaaaa aaaaaaaaaaa	acctccgt			1049

<210> 126
<211> 1626
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (525)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (542)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (562)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (607)
<223> n equals a,t,g, or c

<400> 126

ccacgcgtcc gacgcggcgc	acgcggcagt	cctgatggcc	cggcatgggt	taccgctgct	60
gccccctgctg tcgctcctgg	tcggcgctg	gctcaagcta	ggaaatggac	aggctactag	120
catggtccaa ctgcagggtg	ggagattctt	gatgggaaca	aattctccag	acagcagaga	180
tggtgaaggc	cctgtgcggg	aggcgacagt	gaaaccctt	gccatcgaca	240
caccaacaaa	attttcaggg	attttgcag	ggagaaaaag	tatcggacag	300
gtttggatgg	agctttgtct	ttgaggactt	tgtctctgtat	gagctgagaa	360
ccagccaatg	aagtctgtac	tctggtgct	tccagtggaa	acaaagccac	420

tgcaggtcct ggctctggca tccgagagag actggagcac ccagtgttac aagttagctg	480
gratgacgcc cgtgccta at gtgcylkgsg ggggraaacg actgncccac sggagggaag	540
antggggagt tttccgccc gnaggggggc ttgaarggtc caagttacc ccatgggggg	600
aactggnttc cagccaaacc gcaccaaacct gtggcaggga aagttcccc aaggagacaa	660
agctgaggat ggctccatg gagtctccc agtgaatgct ttccccgccc agaacaacta	720
cgggctctat gacccctgg ggaacgtgtg ggagtggaca gcatcaccgt accaggctgc	780
tgagcaggac atgcgcgtcc tccgggggc atccctggatc gacacagctg atggctctgc	840
caatcaccgg gccgggtca ccaccaggat gggcaacact ccagattcag cctcagacaa	900
cctcggttcc cgctgtgtc cagacgcagg ccggccqcca ggggagctgt aagcagccgg	960
gtggtgacaa ggagaaaagc cttcttagggt cactgtcatt ccctggccat gttgcaaaaca	1020
gchgcaattcc aagctcgaga gcttcagccct cagggaaagaa ctccccccttc cctgtctccc	1080
atccctctgt ggcaggcgcc tctcaccagg gcaggagagg actcagcctc ctgtgttttgc	1140
gagaaggggc ccaatgtgtc ttgacgatgg ctgggggcca ggtgtttctg ttagaggcca	1200
agtattattt acacaggatt gcaaacacac aaacaattgg aacagagcac tctgaaaggc	1260
catttttaa gcattttaaa atctatttctc tccccccttc tccctggatg attcagggaaag	1320
ctgmacatgg tttcctcaag gcagaatttt cctgggttctg ttttctcagc cagttgtgt	1380
ggaaggagaa tgctttctt gtggcctcat ctgtgtttc gtgtccctct gaaggaaact	1440
agtttccact gtgtAACAGG cagacatgtt actattttaa gcacagttca gtcctaaaag	1500
ggtctgggag aaccagatgtt tgacttaggt gaagcattgc attgtggaa tcacaaagca	1560
aatagtactc cagaaagacc ctgtctcaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaa	1620
aaaaaaaa	1626

<210> 127
<211> 1177
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (484)
<223> n equals a,t,g, or c

<400> 127	
ccacgcgtcc gctatcatca gagcatgtca cagatctatg gactcattca tggtgacctg	60
tgttttattt caaacgtcta tgctgcttttgc ttcactgcag ctcttqttcc ttgacgtgc	120
ctcgtgggttgg tggcgtgttgg gttcatccat gcctaccagg tgaagccaca gtggaaagca	180
tatgtatgtt ttttcagagg aaggacaaat gctgcagaaa ttccactgtat ttatatctc	240
tttgctctgtt tttccgttgc atggctttgg ggaggactac acatggccta cagacacttc	300
tggatgttgg ttcttttgtt cattttcaac agtctgcagg gactttatgt ttcatgggtt	360
tatttcattt tacacaacca aatgtgttgc cctatgaagg ccagttacac tggaaatgt	420
aatgggcattt ctggacccttgc cacagccccc ttcacgcccgg ggagtggaaat gcttcctgtc	480
ggangggaaa tcagcaagtc cacccagaat ctcaatcggt ggtatggagg aaggtgccac	540
ctgactggga gagacatcc ttccaaacag gggartcaag gccagccccyg rwttaaagcc	600
aagtccacaa aatgrgrccca cgktcccgcc ctctggagga tatggccagg grtactgtat	660
agccgatgag ggttccagg agtttgc ttaatattt gcattaaaaaa ctgggtgttgg	720
tctcgtgtc agtgataatg aatctggtca aggcagccag gagggggca cccttgactg	780
actccccaga tcgtggagct ccaggaggat acccatcgcc gacactcacc tggcacct	840
cactaaccat tcgactgagc acactttcat attgttatca gcttttgc taaaactctc	900
taagtacatc cacctgtgtt attaggaacct gtgaattgtt ctggatgatt aatacaaacg	960
tgattgttgc atttggatgtt taaattactg attgttatgtt acctgaaaat tcactgtat	1020
aagaaaggtt ggttcgtttt gtatcgttta ataggatgtt catattccaa ggatattgtt	1080
tgtttttttta atcatcctat atggctaaca ttgtttaatgt aaagtaataa tcaataaaagc	1140
aatagaatct aaaaaaaaaaaaaaaa aaaaaaaaaaaaaa	1177

<210> 128
<211> 1276

<212> DNA

<213> Homo sapiens

<400> 128

tcgacccacg cgtccggcca cgcgtccgct	taatatctgt attcccagtt gcctacggga	60
taaaagccca aactccttag cagagaatat aaggccctag	ctcccacatt atttcagcag	120
tcatcaccctt ctatgttctt caagactgca	gccattaact ttttagagtt ccctaaacat	180
gctgtttact ttcatgcctt tatcccggtt	tctgtggaat gacttccctc ttggcccttt	240
tcagtgctac aaacccttat	tcttaagac atagtacaaa tggcatctcc tggttggcat	300
catttcgtca ggcctacagg cctagtaagt	atcttcctcc tctgtgtcc tgcataccctc	360
catttccttg ttatgacatc tataacttta	ataagtacta aaatctgttag tcctacaaa	420
ctcaggcata gaactcattt	cctttatggy tctataatgg aactttaccc aactctcacg	480
ttccccatga ccacagatgt	ggaaaattt aatcttgaca gttcaaggtg aactcagtc	540
ttttcagagt tttcatagtc	ccttcaagat tgaacttcag ttcctgcaat gtttgcct	600
tttctccctt tttgtctatg	ctgggagagg cattgtgggg agggtgtct ggcttatggc	660
tcccattgtc ctctgctga	taaaccacct gagctttgtt cattagcagt ctccctgtgcc	720
tttcacactc aggtagtgtc	tgcacaggcc actctatgtc tttccatgc tgaagaaaatt	780
cctttccagg ccatgtctgt	gttcctccctg ccacacagga aattttttag catgttcac	840
ctccaagctg aatgcagggt	cttgggttagt ggctccacc tgctccagag acttctccag	900
ccattgccac tctccactca	ggtgatgaaat ctggatgagg gactgcaccc accagagtca	960
ggccagggtc ctgtctgctc	tgtgagtccc tccaaattgtt cttatccga gatttccatt	1020
gttctgcccc ctctgactc	ccagggctct caagggagtg ggggtagtga agggagccct	1080
ttcccaagct cccccaagag	ctctagtcac atcacttctg atacttctt tcccaccagc	1140
tggaagaaag aactttcatt	tgtcttgaaa tgagaaaaat gttcttagaa tattttgtat	1200
tactctctgc tctgtcattt	atggtaaaca aaataaaaata ataaaaaaaaaaa aaaaaaaaaaa	1260
aaaaaaaaagg	gcggcc	1276

<210> 129

<211> 1334

<212> DNA

<213> Homo sapiens

<400> 129

tctcagtgtt cagaggctgt	gttggaccca tagagaatt ttccagtcac agacccaagc	60
ttccatgggt ttttactgtt	ctgttaccact tggggkttct gatttgaac ctgtatgttg	120
tgttaattat attttaagca	acacacacac acacacacgc ctcatgtat ggactttat	180
aacaaaagaa aaaatttgg	tttctaattt acaaatggca aattttaat ccctctctgg	240
atgcaccaaa gaccagtaaa	gtttatagtt tttccatcta tatttataaa gcaatactgt	300
attataaaaa tcaatatttt	tatcacatgc ttgaaatttt tattttgttg ttttaaaatg	360
tgcactctaa acatatcaga	accttatttc ttccatgaa cttaagctgc ctgcgcacaa	420
aaaaaaaaaaa aatttacaa	atggagatgc agtagagtcc ataggctcta aaaactaaaa	480
gaaatggat gcagggggaa	caagtttattt gtccctgagtt actgtacttg cttgacatgg	540
ttgttgggtt ctaaatcaca	aaagaatcca ttccaggtat gcatgtctgg ggggtggct	600
gtgtctagat tagaaactgg	gtttcaagct ttgcatgat ggagagcgtc ctctctctta	660
ttagctgcgt gtgttctgga	taggacagta gccggagat gaaaaccacc ttcagtagcca	720
ttagccccacc ataccaagta	acaagtttagg caggaatcgt gggaaatttt tgagtcagct	780
tgagttgttt gagagaatgt	aaacaagatt ggctcgaatt gtaaacgttt gtactttgg	840
tgagttcatg gttctttagg	tcaccttaat accagctatc tttggtagaa gctacagcat	900
ttagttctc tggaaactgt	atcacatttt tgcattttaa aaattttaca gtatcaaaaa	960
accaaaaatct gcttatgaaa	caaaacatga agcaggacat atttggattc tatttattta	1020
aaattaaatt ctttgcaaaa	ttgaacttct caactaaaac gtgtccatgt cagaattttt	1080
actgttagca ggtagttgtt	ggcaaagatg gctaaataat gaagcaaatt agaatctgt	1140
tgtataactaa tgagctgctt	tttttctgtt gagactatca ttatttgc tattacccaa	1200
gaggcaatata	cctgaatttg gatgtctgaa ttataactta tgcaggaata gttctgtaa	1260
tacatttaaa taaactgtaa	agatatttaa taaatatagt atttatacta aaaaaaaaaaa	1320
aaaaaaaaact cgag		1334

<210> 130
<211> 532
<212> DNA
<213> Homo sapiens

<400> 130

ggcacgagcc ttgggccatc tcctaaaaatg atctttatca taatagctac agtaaaaaga	60
aagaaggaga ggtattaatg tgggtggaaa tcaggacagt ttcttaatgc cgtggcttac	120
aattctgaga tttctccagg catcaggaca tgtgcgcga caggacttgg ctcttttagg	180
agatacttca gtttgtatca gatgtggctg tggagggtgc ttcttaagca ttgctaacta	240
tgagtgggtc ccttcagaa ggaaggactg taagaggtat gaaacctctg agaaaacgag	300
ctgtcttc ttaccaagcg cctgcagccg tcaaaatgct gttaggctta gtcgtctgcc	360
agttcccaag ctgagctgtc tccttcatgg ataggatttgc tttgtttaga aacaacaaca	420
aagttcattc tgtttataac tcagagcatt tgtttttct gctgaggcta aaatacttgt	480
ttatttttc cttagaggaga aaagaaaaaaaaaaaaaaaaaa aa	532

<210> 131
<211> 685
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (491)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (661)
<223> n equals a,t,g, or c

<400> 131

tcgctctctt ctctctctt ctctctctg tttaagtc aagtatttgt caaaaaaaaaatg	60
caatcttcgtt ttttttgc acgacaaat cattttcttc gtaagcacct ttttctctcc	120
actctgtcac tgcctgtgtg ggtaactgggtt ataaatgtgg aaaaagaata gttatgactg	180
taacagattt ttattttat ttcaaaattt tatatgaatt atgtatatct taatgatcg	240
tcattttccc agttgtaat atatgtgttag aaattgcctg tatatgatata tgcttttct	300
cctctccccc tctctttctc tcctctccct ctctctgtct ttctccccgc tcaactgtct	360
cttttctttt tgggttctc ctcccactcg gtgcctctgg tgcgacttg gcagtcaagg	420
agaggcatgg tggcctgggt taggaagagg gaccctgtcg cttagaaaaag cggagagtga	480
gattgttagta ntcttatgca aaagctattt ccagtatttc ttagcagctt cagaggtatc	540
tctcaactccc tgcggcgc ttactgttt atcttaact gctgttttat ctatatgtaa	600
aaactttcta aagcaaatac agtatttc tttttcttat caaaaaaaaaaaaaaaaaac	660
ncgagggggg gccgtaccat tcgccc	685

<210> 132
<211> 729
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (725)
<223> n equals a,t,g, or c

<400> 132

tcgacccacg	cgtccggcca	tttagaaaata	atcaactctt	aatcagcctg	ggatagtca	60
tactaaaagc	acccatcatga	gctgtaaaaa	attnaatgca	tttatttaca	tattnatgtt	120
taaattttag	tatattgtta	gttgaggat	agtttccaaa	caaagagccg	tgaaatgtt	180
agtaactgtc	tctgtaccc	ttggatgagga	cagctcagcc	gggaatggag	ggggactggg	240
tgaggagacc	agaatgtcag	tgtggccacg	cagcacactt	ttgtttgtc	ttctgtcctt	300
gagcaactggc	ttgttcctgg	ataaaactagg	cataataata	cctatcctgc	tgtgtgggt	360
gaagttaaat	gtgataatga	tgtgtgtgag	atgcctgcac	agtgcctgga	ggtattgaag	420
aattattgtct	gcctwttctt	tttctaccta	ccacttaccc	gctacccccg	ggtgctacat	480
gttagaaaaac	actgtgtaaa	gtgtggatgc	ttctgaaaaaa	tctcoctgcc	agcagttgt	540
gccaatagcg	tgcagaaaaat	aagatgcaat	gatttggctt	cttttctgtt	tggcaataag	600
aagcttattt	gcmcatagcc	tgatttctt	caatctgca	aaaaaaaaaa	aaaaaaaaaa	660
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	720
rgggngggc						729

<210> 133

<211> 1079

<212> DNA

<213> Homo sapiens

<400> 133

ggcacgaggt	gttagaaagt	tttcgaagca	gtgtgagtct	tgtacctttg	tggcctgtc	60
tcacagacac	ctgtctattc	cctgaccctt	ttaaatgcta	actttctgcc	tgttagaaat	120
cttccctttg	tgcttaggtc	tttttcttct	gtgagctta	gataaacaac	ctagtgttta	180
aacttttaa	taagggattc	attttttaat	acatgagaat	tcatttcaaa	attttggttt	240
tagttattta	ttttattctt	cttggctctt	tttcagacag	atgttctctc	ctggattgta	300
aaagtgcata	tcaaaggatt	tttatttgc	atataacttaa	cctttctctt	gtaagttgcc	360
atctgtgtag	atacagcttt	gattgctga	caagagggaaa	atgtttccc	tatcttttc	420
ctgcctgaac	tatacggtca	cttgcgttcc	agcatagtg	ttcttaaccc	tcatagtgt	480
tcagaatcac	tttgcagagc	ttttaaaaac	tctagatgccc	tggggaccac	cccaaagact	540
ccatTTTGT	gtcatgggtc	aaagcacagt	cttctagttt	gcagctagtg	ttgagttacaa	600
ctagagttt	accagggttga	attttagttt	aatcttggct	ggctctgaag	atgttagttaa	660
tctctattca	tttttttkga	aaagttaccaa	tgaratcaga	aagtttattt	gaaaacatct	720
agttgaatcc	cctgtttta	atagatgggg	aaaccaagac	ccagagaata	taatccaaag	780
ctacctgtca	cataggccac	aatttctttt	ccaatattct	gttctcgct	gttcttctaa	840
tttgcagaac	tcctctttt	aaaacctttt	gagaatgtat	tggcctcata	cccttcttct	900
tcagcctgaa	agacatgcac	ctgtcactt	tttatgtat	ttaaatgca	cctctagaac	960
agggggtgtcc	aatcttctgg	tttccctggg	ccacattgg	agaagaaatg	tcctggccaa	1020
cacataaaat	acactaatga	tagccgatga	actttttttt	aaaaaaaaaa	aaactcgta	1079

<210> 134

<211> 1297

<212> DNA

<213> Homo sapiens

<400> 134

gactcgtgcc	gaattcggca	cgagggssaa	gggcgtcttc	tgtcccttgt	gtccgcttct	60
aggctttttt	gagcatctt	cagggtcact	ccagagaccc	ccccagacta	ctgaatctga	120
atccgtttag	tggtccaacc	ttccctttca	ggtgtctact	agagggcaagg	ttggatgtat	180
gcggctgccc	tcagtagcgc	cccttcctt	ttttttcttc	atttgtgtt	actttttttt	240
ttaatcctt	tctccctctc	ctccatccct	ctccctcccc	tactatacag	ttatgactta	300
cactaatcat	tcccgtat	ctcccagaag	gaaataagt	tctgtctgtc	tctgtctctg	360
taccatcg	ttcttggta	agttcgact	atttttctc	tacacccag	ttatgcaaaag	420
tgttagtccct	gaaccagcag	cctcagcata	ctgaggaaat	ttgtttaaaaa	cgcgaaatct	480
caagccctac	ccagamttac	tgaatcgt	tctgcattgc	aacaagatcc	cctggtaatt	540
cmtatgcaca	tcaaaattt	gaaggcaca	ctctcaact	atgtcttggg	tctcccttcac	600

atccatcctg ggaaggtctt atccctcatt cctgagctca tcccactgaa	660
acttccaatt cctgaggcctt tgtgaggttc tgctgtctag taagttgtct	720
acctccgaaa acacttgggt ttcagtttc tctgtgaggc ttcttaagga	780
gtggatgttt tcaagataac gcagctaaca ttcaaagagg ttaagtgaat	840
cacacagcaa gcactggaga ggcagtgtct caccatgttg cccggactgg	900
stgggctcg gccatcctcc cgtctcgcc ttccaagtgc tgagatcgca	960
accacgtccc accggatac ataggtttgc cggtatcctc tgaacctccc	1020
agagtggaca aaactgtggg tcccycmtyt tcaaaatggc cagtaaaaga	1080
atatgcattt ttagttatcc tctgctgccc tcttaagtt gattggggat	1140
ctactttggg aagataactt accttcttat ccactatggc taattggagc	1200
tcttatggt tgctggaaa tttcaataaa aatttcactg ggaatggttt	1260
aaaaaaaaaaaa aaaaaaaaaaaa aaaaaatgac cctcgta	1297

<210> 135

<211> 617

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (9)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (513)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (559)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (587)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (602)

<223> n equals a,t,g, or c

<400> 135

atggaaaanc aggcaaattcc tgaaatgggc tggaaaaaaag ggagggaccc	60
gggagaaaac ttggcattyc ttgggaatct aacaggatgc agtgaaccca	120
agagctcacc aatcagactg cccttgcata tccatgagca gatgttgtat	180
aggccctcta gtgggtatgc tgccaagcaa ctggagtgcc acttgggctc	240
gtctatccct ttcaccctgg catttcatca gccaaacaaa aaccaactaa	300
aaggaaagcc cctcaagggt cctttgaccc cgatatctac atagatgcta	360
ctgaggggta ccaaacraat tcaaagctcg aaatcaaata gctgctggat	420
ccttttcttg tggctacta taaataaaaa tggtagactgg ataaattaca	480
aaaaaaaaaaaa aaaaaaaaaa ctcgaggggg ggnccggta ccaattcggc	540
tctgttattaca atcattggnc gtcgtttac aaatcgta ctggggnaaa	600
anccaattta atcggt	617

<210> 136
<211> 1311
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1284)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1296)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1301)
<223> n equals a,t,g, or c

<400> 136

ggcacagctt ttcaacatgc	cittcagcact aatgactgct	ccaggaatgt ctacattaag	60
aagaatggct ttactttaca	tcgaaacccc attgctcaga	gcactgatgg tgcaaggacc	120
aagattggtt tcagtggagg	ccgccccatgca	tgggaagtgt ggtgggaggg	180
actgtggcag tgattggaat	tgccacaaaa cggggccccc	tgcagtgcca aggttatgtg	240
gcattgctgg gcagtgtat	ccagagctgg ggcttggaaatc	tggtggacaa taatctacta	300
cataatggag aagtcaatgg	cagttttcca cagtgcacaca	acgcacccaaa atatcagata	360
ggagaaaagaa ttcgagtcat	cttggacatg gaagataaga	cttttagctt tgaacgtgga	420
tatgagttcc tgggggttgc	tttttagagga cttccaaagg	tctgcttata cccagcagtt	480
tctgctgtat atggcaacac	agaagtgtact ttggtttacc	ttggaaaacc tttggacgga	540
tgacagtggc tttttgtga	tgacagacas aatggaggag	agatctgctt atgggaagta	600
saaccatgaa gtgactgtca	cacatgcatg tccaagaaaac	atcctgaaaa cacatgaagt	660
cgtaaaactgg agaagcagct	ctacagcaga gattatctcg	tgtttccctt ttctactggg	720
ccagaaaaat cctcagggtt	gcagttgggt gagtgggca	ttgacatatg catgttgac	780
ccgatgttgt ctctaagttt	gcaatgtgtt atttccagct	ttaaaggta gattgttagag	840
atgctgtcaa aggataagg	aaatagcaag atttttaagt	agtgtgtttt tgaagactga	900
tcccattttt caactgcctg	ttctttctcc agtccttttt	tttccagcca gcttgactat	960
tagaaaaagta taaaactgg	tgggttttat ttaatatttt	taatattttt agaagcatgg	1020
tctgccttga ctgcacttct	ctaaaagtga gatataaaaat	tgtgcagcta tttaaaaagt	1080
tgtatataat atgtgtgtaa	aaaaaaaaac taaaaggaa	raaggacaaa caggttgg	1140
tgttctagtt ctaatttctt	aaaaaccact acatggttac	aaaattggaa taacattttg	1200
gggggacaac tgggttaac	taccaaagaa ggagggattt	aaagaggaga tgggtgttga	1260
attgacccca ttttggataaa	tttnaggctt acagtnccca	nagctgttag a	1311

<210> 137
<211> 1095
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (616)
<223> n equals a,t,g, or c

<400> 137

gatggtatgt gtgtgggtgt	tatagggtga atgtgtgg	tgtgaggtgg gcatgggtgt	60
tgtgaggtgt gtgtggatgt	tgtggcatgt gtttgggtgt	tatggaaata tactatgt	120

taggacatgt gggatttca aagatctatc cttttgtgct ttgaaatctg	180
actgtggcct cactgaggag gagttttaga atatgcagg gagatgatca ggactggatc	240
ttgtatggc gtaccacatc cagtcggaga cagcatgcta aggcaaggag ctcataaaag	300
ccccaaagctc tagctgtgg ctacttatct cctggagcat caggtgagcg cggttcaggct	360
ggggagtcct gatggctgcc tggttttac agatgttac agcttaggccc tggggacata	420
gcccaagcacc ctccagargt tgggtctgtt cttaactctt caggttcctt ggaggcagga	480
gaggargtgg cctcatttctt ggcaggcacc ccactactgt tattgagcaa tcctccaggc	540
tgcagagatg tcagaggagg accctaattgt ctcckgattt tgattttt gttcttttc	600
cctaggtgtt ttactngcag ataccttgag taccttgttt gtatattcac tttgaaagca	660
cacatttaaa tggttataag gaaaagggttc taagacatc cattgatcca ttcatttcatc	720
attcagcaaa tacctgttga atacctgtg tggcttaggc actgcgggtgg gcgagccaga	780
rggctttgtt gctccaagga rcttgcattt tagtatttcta gttattttca cgcattctgca	840
ctatctggga cagggaccat tgcgtttgtt cgtatataaa gcagcatgtg tctgcactac	900
agtttgtgtc cgytgcagat gggcaaggat tgagtgcataa aacttctggg ccaaaagggg	960
ttggcttggg tcaggctgtc aagtagctga ggtgaaagca tggccaccc ctccgtatac	1020
agggatccctt gctgattgtg tggacacca gggccttccc atctgtcagc tgggtttgtc	1080
ctcacagtag ctcga	1095

<210> 138

<211> 692

<212> DNA

<213> Homo sapiens

<400> 138

ggcacgaggc gaatatgtgt agtcagctg ttttggaaat gatctgtttt tagaaggcca	60
caaaggcaaat attattatct taatcttattt ctgaattttc accactaaaa ccacattcta	120
ttgaaggaat atataataaa agtgcattat catatgtgt cacaatgagg gattcagggt	180
cgaagggaag actcattctt gtgaaaacat agcccatccc cagcagttgg tagaaggatt	240
tgcgtggagtt cctccctttt gtgtggccta taaaacattt catgaggcat gtggcaata	300
tcacaatgt agtggcttta ttcctccag tcttagcatc ctcactcaag ccacctctt	360
tcatagacac atacttatgt ttgggaaga ggtgctctag gtgggacacc ctcgtctgt	420
ccaaataatt cctactgaca tccatggcag cttcatttcta tctgagctgg agatttggg	480
attttaggtgg gcacagaaga aagaagggtt ttggggcagt gtcgtttgga tgattttgac	540
agattcttcc tggggtaaa gagagatagg tgggtctaat catccaggaa ataaaatgcm	600
aagggtgtgtg tataatggaaa atccaaggaa gaggaaatta aaattatccc agattgctt	660
ttaatagtc agaaaactca actttccatg aa	692

<210> 139

<211> 748

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (60)

<223> n equals a,t,g, or c

<400> 139

ggcacgaggtaagggtttccacagga cgggacaagg tcttagatgt ttctcaatan	60
ataaggaatg aatctctggg ttggccaaatc ccgaactcat tagctctgaa ctcccaacac	120
acattcaggtaatcgttgcatacaggatc ttctcagggt atgctcaagt tattgctatc	180
gggcacatcttccatcaga attccagcag aaatacccaa tgggagtggt gggctctggaa	240
acaggaatgt gggcagagct gaagctgtc tccctggggaa gggctgtat tgctgtgtgg	300
gtgtgcctgtcga gaagagtgt taggggrgga cacagtccac cagcaggtca aggtggcag	360
ggagtttaagg tccagtggga aggagtgcag ggatcaggaa gtggccagcc agaagacatg	420
agatgggaga agctacatgt gaggattctg atgcaggcata tgcacaggat	480

gacatcagat ctgtcccg	ctccacagca tttcctgact	gcctcatct accctgcaga	540
cccacctgcc	ctggggtttc	cttggatct ggctgaccag	600
aggctggagg	agagtggatg	gtgagaccca ggctccagc	660
gtggtcaggg	ctatcaggtt	tctcacccctg ccaggccaca	
tggcgggtag	ctaagccaa	actgaggtcc aaggggagtt	720
	cactcgag	ggtggcagg	748

<210> 140
<211> 1132
<212> DNA
<213> Homo sapiens

<400> 140	gacacgagca gctaccctta	tttgcagaca tcctgagaat	gcaatatgat gagcaacttc	60
	tgtttccccca	atccttgctg aaggtaaaata	agatttattt tgccaaaac tttcaatct	120
	acaaaatatt	tttttctta taaaatgctt	tttgtctt gttcatgtc tatatattt	180
	ctaagccctc	tttttctccc attgcatggc	tctcctcacc ctcacttta cttttgttt	240
	gctgtttgca	ggacctcctg	gtctctgtca gaaaagactt	300
	ttgcaacttc	ctattttctt gaaatctatt	tactaaatgt tatgcaagac atgaaatttt	360
	tgcctacttt	attccaccat catttgtttt	atgaaacaaa caaacaaaaa aatctctaaa	420
	cctaacccaa	gtagaagatg	cttaacttta agaaaacact aggcaacagg	480
	gggtgtcaca	ggatatgaac	cgttatggaa ctggaggagg	540
	aactagcagt	gtctctccca	ttctctcaggg cccaaccac	600
	ttctctctat	ctcctctttt	atggtttcca gtttcttga	660
	ccttgaaggc	cttttcatac	catggatttag ttaacgaact	720
	aggatgtatc	actggatagc	ttctctttaa	tagaacatga
	aaggtaggca	taatttggcca	attacctgtc	780
	aggcagctgc	acaggcctct	catgtgtatgt	840
	acaggcctc	cccatctcag	gattaattaa	900
	tagtaaaaaga	aactggggtc	ttgtgtgtg	960
	tgtgggggtt	agcaactcta	ttgtgtgtg	1020
	tatattcaca	aatatctgta	tatttgatta	1080
	agggctctct	atattctgct	gaaaacccac	
	cactacccta	aaattttttt	agagagatca	1132
	tcctccaatg	tcataactgc	ataatgtgcc	
			cccacctact	
			cacttccatc	
			tg	

<210> 141
<211> 1112
<212> DNA
<213> Homo sapiens

<400> 141	gtggcaaatg	gggcctactcc	tgcttttact	tcttgggtgc	tggaccata	tattttttac	60
	aatggaaatg	atttatttggt	atcttgaagg	ccaccccccatt	cttaatgaga	tcctcttcat	120
	tctgcacttt	taaagggtat	tgcagcaactt	tatcaggcca	gcagctttgg	gtaataactg	180
	tatgtggtag	gaactgtgga	tcttgggtt	catatgtgt	cattgtacct	cctattctgc	240
	aaagtggaaat	ccttggttcta	agatactatg	tgagtttct	tgttagtcaa	tgagataactc	300
	ctatgagtcc	ttggatagta	atactggtca	aggcaccata	gacagcatag	gaaaatgtat	360
	atagtagtcc	ccctttatgc	atgattttac	tttctgtgtt	ttcaattacc	tatggtcaac	420
	cttggtccaa	aaatgttaaa	ttggaaaattc	cagaataaaa	taattcataa	tttttttttt	480
	gtttttgaga	cagggtcttg	ctccagccca	ggctggagtg	caactgggtt	gtcatagctc	540
	actgttagct	ccaaactccctg	ggcttaagca	gtcaaggctc	agcctcccaa	atagaactac	600
	aggcatgtgc	caccatacct	ggcttagtctt	ggctattttt	attttttattt	ttataggat	660
	gcagtcttgc	tatgttgc	aggctggtct	tgaatttctg	gcctcaaata	atcctccac	720
	ctggcctccc	aaagtgtgg	gattataggc	aaaaactact	gcacccggcc	ctaattcata	780
	gtttttaaat	ggtgcactgt	tttgagtagc	acaatgaaac	tttgagctgt	cccatccat	840
	ctggcctggg	atgtgaatca	ttttctgttt	agtggctcca	caactgtat	gttacccccc	900
	cattagtcac	ttaatagcca	tcttgctt	cagaccaact	gtgagagtt	tgctgtgtt	960
	atgttaagta	acccttatat	tacttaatga	ttatccaaag	cacgacagta	gtgtatgtgg	1020
	caattcagat	gtgccaaaga	gaaactgtaa	agtgcctcct	ttaagtgaaa	aggtaaaagc	1080

tctcaacaaa aaaaaaaaaaa aadaactcgt ag

1112

<210> 142
<211> 1084
<212> DNA
<213> Homo sapiens

<400> 142

gtttgggg	catcacagac	tacaccgta	ttagaggatg	aacttaatg	ataaatttg	60
tgtgtgtgca	tgcgtgtgt	cgtgcgtgt	gactgttaca	ctcattggc	cttctgtgt	120
ctctctccct	ctcctcagcc	ctcttattc	cctgggacac	agaaatttt	aaataagcc	180
aattaataat	cctacattgg	tctttacgt	tttagagtga	aaagaagatt	cacatatctc	240
tcattttaaa	ttgaaagcta	gaaatgatta	agcttagtga	ggaagccatg	ttgaaagctg	300
agatagtcca	aaaacttaggc	ctcttgccacc	agtttagccaa	gttgtgaatg	caaaagaaaa	360
gtgcctggag	gatattttaa	atgctgtcc	agtgaacaca	caaacgatag	gaaagcaaaa	420
tagccttatt	gctgatatgg	agaaagttt	aatggtctgg	atagaagatc	aaaccaactg	480
caacatttcc	ttaagcaaaa	tcctaattca	gaacacagcc	atagctgtct	ccaattctat	540
gaagacagag	cagagaggaa	gctgtgaag	taaagtttga	aaataagagg	ttgttcatga	600
ggtataagga	aagaagacat	ctccataaca	taaaagtgtt	agtgaaacat	caagtgcgaa	660
lacayaagct	gcagcaagtt	atccagaaaa	tctaagatca	ttgaagaagg	tggctacact	720
aaacaataga	tttcaatat	agacaaaaa	gccttctgtt	gattttaggc	atctagctta	780
aatggaaag	agatgccatc	taggactta	atgggttagag	aggagaagtt	gatacctgtc	840
ttcaaagtta	agactgactc	ttttgttagg	ggctgttgca	gctggtgaca	ttaagttgaa	900
gccaatgtc	attcaccatt	ccagaaatcc	ttgtgcctt	aagaattatg	ctaaatctac	960
tctgactgt	ttctacaagt	agaacaacaa	agcctggatg	acagcatatc	tgtttatagt	1020
catggtttac	taaatatttt	aagcccactg	ttgagaccta	ctgctcagaa	aaaaaaactc	1080
gttag						1084

<210> 143
<211> 1050
<212> DNA
<213> Homo sapiens

<400> 143

ggcacgagtt	tatgggtgc	tgaccactcc	cactttcaca	gaaccctcat	60
caaacagcct	tctatgatcc	caaatacgtac	tttctatcac	attttatgc	120
cctactcatg	aaaatgttgg	ggccatccag	gcttccattt	ttagccctca	180
gtttatactt	tatttcagt	tttgttatct	gatctctgac	tccagccccag	240
actccacatc	cacatattca	tctggcttgc	tgaataactt	ctcttggatg	300
ccttagactc	attatgttca	gacatgaagt	catctttttt	ctctccagac	360
tctcgatattc	ttctttttgg	tgaatggtac	aattattttag	atggAACGTC	420
gtcggtctag	aatcctccct	cactcctaat	gcacatcca	aaatcctatc	480
gattcggcct	tctaaataca	gtcaaaacat	ttcattcaat	tcagcgtcac	540
ttaatgtaga	ccttctctat	tttaccatga	tcaagcagag	gcccgttac	600
tgcccttccag	tcttgcatac	ctactccgca	gttaatcccc	tgagtgttat	660
cttcttaacag	tacagattt	gtcatggatt	ctccagttt	cctagtgtat	720
ggaaacatgg	atggagatgg	aggctattat	acttagcaaa	caaatgcac	780
aaataccaca	tgttctact	tataagtggg	agctaaatgc	tgacaactca	840
caaatgaaca	gcaaacactg	gggtctactt	gagggtggag	tttgggagga	900
cagaaaaggt	aactattggg	tactgaactt	aatacctggg	tgattaaata	960
caggccccca	tgatatgagt	ttacctacgt	aacaaacctt	cacatgtatc	1020
aaataaaaagt	aaaaaaaaaa	aaaaaaaaaa			1050

<210> 144
<211> 1113

<212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (349)
 <223> n equals a,t,g, or c

<400> 144

gttgtgttg	agcacagctt	taggcttaga	ttcttcatca	actaggagaa	gctgtgcttc	60
aatacaggtt	ttcgtttgca	tggttcctaa	tgtgcttcac	tcaattttagc	agaatttttt	120
ttttaaccc	ttccttgacg	ctagctgctt	gtgcaaatac	catctggcc	gcctactctt	180
cttcaactgc	tgacagatgt	gtaggtgaga	aaagtctcat	agtcttgc	cctgaaaagaa	240
gcttccagac	ccacttctag	ggccagtgtac	atatgcagga	aatcagctgc	ttctgggcca	300
ggacagagct	ggttttttt	tttagtggggg	atggcgggca	gtggggcang	ggacattcaa	360
aatttatttt	ccaacagaca	gatagcatca	gcaggtacaa	ctacaagggt	atctacatag	420
atcatacatt	cacaaggcat	tattagttca	acagtggaaa	agccactcgt	gggtttctg	480
taacaatatac	ccacttcata	gtgttaaacag	gtactatttt	gttcaactac	aattccggaa	540
ggaaggggcac	accctgcagg	ggggaaagaaa	aggggaatcc	taaagtaagg	tgcaacaatt	600
aagagacaac	actttggcta	acaatcttgg	atccacattt	cagtcaggc	cttccacata	660
gagggggaaag	acttttctc	cagaaggtag	aatctttctt	cctcccttct	tgttaaactg	720
agagcagtgt	tttggggct	caatattaca	tgtacaaaag	gagattagaa	gaaaatgcac	780
cacaacca	tcttgaacgt	tcaagcttcc	ctgccaatac	atcacaactc	ttaggtttta	840
gacggggcct	gggaaatacgt	aagtgtttt	tctttttttt	tttttaagt	gaaagcaagt	900
tttattacgaa	agcaaaggga	taaaagaatg	gctgctccat	aggcagagag	cagcccagta	960
atcttaaaat	aggaaaatag	acactatggc	tacaaaaaaat	aaaaaataaa	tgaggttagat	1020
aaaattttca	caccaggac	ttgcctgttc	caacttcata	gtcttcatga	aatattcatc	1080
aagaagacaa	aaaaaaaaaa	aaaaaaaccc	gtta			1113

<210> 145
 <211> 685
 <212> DNA
 <213> Homo sapiens

<400> 145

ggcacgagca	cttcctgaaa	taaaaggggag	ccgcttacaa	gaaataaaatg	atgtatgtgc	60
aatctgctat	catgagttt	caacatctgc	tcttattaca	ccgtgttaatc	attatttcca	120
tgcactttgc	tttcggaaat	ggctgtacat	tcaagatact	tgtccaatgt	gccatcagaa	180
agtatacatac	gaagatgata	tcaaggataa	ttcaaatgtt	tcttacaaaca	atggatttat	240
tccacccaaat	gaaactccag	aggaagctgt	aagagaagct	gctgctgaat	ctgacaggga	300
attgaacgaa	gatgacagta	cagattgtt	tgtatgtt	caaagagaaa	gaaatggagt	360
gattcagcac	acaggcgcag	cagctggaa	aattttatgt	tgataactgac	tgatgaaaat	420
agcatttttt	aatgatttgag	gtattttttt	aaaattttagt	tcatccaaaa	tggagtaata	480
tccttcaccc	tcagttgtt	accaaggaca	aaaacagttat	caatgttcaa	tctgtgaatg	540
gttttccgtt	tactgtgtt	tgctactgtt	aatatacctc	ttttaattact	tctggctctt	600
ttggtgaccc	gtttaaattt	gtgtacatta	ttgtacata	aaaaaatgtt	tttcacattt	660
ttatgacaaaa	aaaaaaaaaa	aaaaaa				685

<210> 146
 <211> 1038
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (743)

<223> n equals a,t,g, or c

<400> 146

ggcacagtga	agccatctat	tgacaaaatgg	aggagaatga	tgaaaagttg	taagatgtt	60
aaatattttg	ctattattca	tgattattcc	taaattttat	ctttcaaac	tgttgctact	120
acttcagaag	attacacatt	ttatctgtgg	caagacactg	aacaatttaa	attttaggtg	180
tgaatcatat	ttccctgttc	tgtacctcta	tttgtcatac	atattatact	aattcttaac	240
aggaaaaaat	actttttck	tttttatctt	tgtactttt	tttgaaggtt	ttaatttgtt	300
ttagagatgg	aataaaaaag	ctgtgatgat	atataatcct	aataataaaa	ttacttgatc	360
aacggtttg	aaaaatacc	ttwaaaata	atagggcatg	gtggctata	cctgtaatct	420
cgtactttgg	gagggcgaag	tgggcggatc	acctgaggct	ggtctgaac	ttctgrgctc	480
agacaatctg	tccacctcg	cctcccaaag	tgcgggatt	acaggcatga	gcccaactgca	540
cctggctgtt	atgtcccc	gatgaatcaa	ctcttcata	attataaaatg	actttttta	600
ccccctggtaa	tgtccttgc	atagaaatct	ttttttttt	tttaaataag	aaacagagtt	660
ttgctctgtc	aggctggagt	gcagtggtgc	attataact	aactgcagtc	ttgaactctt	720
ggccacaaatg	gatccctcctg	ccnyggctca	tacctgtaat	cccagcactt	ttggaggccg	780
aggtgcgcgg	attgtctgaa	gtcaggagtt	tgagaccagc	ctggccaaca	ttgtgaaacc	840
ccatctctac	taataataca	aaaatttagct	gggcatggtg	gtgggcacct	gtaatcccag	900
ctactcagga	ggctgaggca	ggagaattgc	tgcAACCCGG	gaggcggagg	ttgcagttag	960
ccgagatcac	accactgcac	tccagcctag	gtgacagagt	gagactctgt	ctccaaaaaaaa	1020
aaaaaaaaaaa	aaactcga	g				1038

<210> 147

<211> 851

<212> DNA

<213> Homo sapiens

<400> 147

ggcacgagaa	caaattgata	gtgagcatta	agggtttcca	agttggattt	gtaactcctc	60
atcattcctt	gtatgacaac	tttctgaata	tatgtcacta	tgttagaaaa	ttaaacactc	120
caaactcata	tttctgtgt	tagaagttt	cagcggtact	tccatgcaac	tttaaatctc	180
actgctctct	atgggtgatg	tcaaattgacc	ttcagtaatg	actgagaatt	gaatacaaat	240
agattacaaa	gccaaaattt	gatgttaaat	gactcaggaa	attttagttg	tatTTcaat	300
tcaagtactt	agtagcctac	gtttgcttgg	cctctgggtc	tttatggaaa	ataggcttg	360
tagtggcatt	gtggagcaaa	ggagactgtt	acaccttaat	taacttttt	tactgatgca	420
aataatttga	ggatagagag	gagggaaagta	gtgaaagcta	tgacctaaaa	cattgggacc	480
aaatagagc	tcacagatat	ttggattatt	ttatgtgctt	attattaaat	aaggaaagca	540
ttttgtgata	tgtgaaagac	gctatgtgaa	gttttaccta	tcttctcaaa	gacctttct	600
tttgtatttt	cttttgggt	ttcttaaagc	caaacaaaaga	aatgtctta	aggagacagg	660
gtggggtttt	ctgtgggcct	ttgttgggtt	ttctgtkggc	catcccctc	taatgaaatt	720
gatctctggc	tgtttgattt	ttttcatatt	gtatTTTAA	aatTTGTTGT	acagtgcct	780
gtgagcacca	agtaccacta	gatgaataaa	acgtattata	tctaaaaaaaa	aaaaaaaaaaa	840
aaaaactcga	g					851

<210> 148

<211> 614

<212> DNA

<213> Homo sapiens

<400> 148

ggcacgaggc	aatatccact	ctacccagct	ggggccccag	tctacaaccc	tgcagctcct	60
cctccctata	tgccaccaca	gccctcttac	ccgggagcct	gaggaaccag	ccatgtctct	120
gtggccctt	cagtgtatgcc	aacttggga	gatgcctca	tcctgtacct	gcatctggtc	180
ctgggggtgg	caggagtctt	ccagccacca	ggcccagac	caagccaagc	cctggccct	240
actggggaca	gagccccagg	gaagtggaaac	aggagctgaa	ctagaactat	gaggggttgg	300
ggggaggggct	tggaattatg	ggctattttt	actgggggca	agggagggag	atgacagct	360

```

gggtcacagt gcctgcatttc aaatagtcgg tctgctccca agatccccaggc caggaaaggc 420
tggggcccta atgtttgtcc cctctgggt ggggtggggg gaggggaggag gttccgtcag 480
gcagctggca gttagccctcc tctctggctg cccccattggc cacatctctg gcctgttaga 540
ttaaagctgt aaagacataa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 600
aaaaaaaaaa aaaa 614

```

<210> 149
<211> 1200
<212> DNA
<213> Homo sapien

<400>	149					
ggcacgagga	gagagaagat	gataaaaaga	ctgttgatga	ccgaactgtg	aaattctccc	60
cttgcaccc	ggaagatggc	atgggcctt	ctgtccgtct	tctttcttcg	ggcttttgtt	120
gctcaacta	gcacagcata	caagtgtgt	ctttgttcgc	ccaggtctcc	atggtagtt	180
gaagccaatt	tctggcttga	cttttatggg	aaaagttatt	ttatgtctcc	taagcattag	240
agttttcta	ttactctatg	tagttgagac	aggatttgat	aagtcttaga	aaagaaaagat	300
gggaaaacgg	gattcccttt	cagaagtacc	tgtgtgtatc	tgttaataac	cacaggggtt	360
aatatgatgt	aggatctttt	actatcaatt	tcaaccattt	gatTTTGTAT	gattgaaact	420
tgcaccgagc	tttgactgtt	tgttaaagag	tcatttttaa	tgaagaata	attctttatt	480
gctggttttt	catttacact	gataaataca	cagatcttaa	taaagtcttt	aacattcatt	540
tgtattcaga	tgtgagtaga	agaactaaaa	aaagaaaagtt	acatatcact	atgactgaag	600
gtacttcagc	ttaatctgaa	atataattta	acttgtgaac	tccttggata	tgatattatt	660
tggaataaac	agaatttatac	attgaaccca	aagttagggaa	tgatagctt	cattgtctaa	720
aaatccttac	aaggttaaga	tgattcaata	tcaagaatat	tcagaaaatt	atttctaaag	780
ttgatcgatt	catgtcgat	tgatagaatc	ttgaccagaa	gaaattttgc	tcttttata	840
tagttcaag	aaatgtgttt	ttaaattttt	attaatgcac	ttgaacaact	ttgcaggaat	900
aaagcaaccc	cctaaccaca	aaatatccct	ctaaatttagt	tccctagctt	tctcaatgaa	960
tacacacata	ttttcacata	gctatgatcg	ttgtgtacat	tctcccttgc	tttacttctc	1020
ggcctaacac	ttgtctccct	ttgtcaacac	agattctact	ctcaccattt	taaatgtctt	1080
tatatccatg	ttacatgggt	aacctcaactt	caccccattt	ttagatattt	gagttatatac	1140
taattttca	ctcttataaa	taqtqctqct	atqaatqct	gtaaaaaaaaa	aaaaaaaaaa	1200

<210> 150
<211> 683
<212> DNA
<213> *Homo sapiens*

```
<220>
<221> SITE
<222> (41)
<223> n equals a.t.q. or c
```

<400> 150
gggagaatag gttagaagaa aatatctacc ctgagacagg naaaaacaag aaaatgacat 60
atacagaaaa gacatatacg gacacaggaa ctggtagtaa agtctatgaa aggctgttgt 120
ctctcaccta gtgttgcc taaagtgcgt gcagtcagaa aggctgcag ttaaaaaggaa 180
aagttggatg tcaagtggaa aatactgaga agaaaagttc tcagaggaag caatatcctc 240
cctaaaaacc aggaactttc gaacaactca agggtccact ttcaggtcaa gtgctgaaag 300
gcaatgcaga tgacagttgt atggtatgtg attactgcaa tcatttggtg gagaatgagc 360
atgtgtgaag ccctctcaca gaattgcttc taatcctaaa atgtatctca ctgtgtatgaa 420
aaacaatcaa agtacagttt agactaaggg atgtgtcctc aagtttagca gactctaaa 480
cacttacttc tagttgttagc ttgattattt cattttgttt ttctttttct tgacttctta 540
gctttgcatt taactctgaa atttccatct cctttttctc tatttagttct ttgtgtttt 600
cttcatttaa ttcaactgaa taaaatgaaa taaataaaaat tcatttggta aaaatttcaa 660
aaaaaaaaaaa aaaaaactcg tag 683

<210> 151
<211> 827
<212> DNA
<213> Homo sapiens

<400> 151

gggcacgagc ttgggcctca	agtgattctc ctgcctcag cctctcagga	caaccccagt	60
tctgtcatcc acgttgtcaa	tcagaccaat gcccaaggcc	agcaagagat	120
tgctgtctcg aagcggcacg	agcctcccccc	agccctgag ccacccctcag	180
ggaaaagctt caaggaatag	ctgaggagcc	agagatccag atggtttga	240
ccagaccatt tcttccccag	gtcctgaagt ttgagccagg	caagtggcag tgcccttagt	300
gggcagccgt tgccaatgga	tgcctttagg	agtggtgccg agagcagtgt	360
ggcctgggtt tgcatttc	tgcagactct	aaagacttcc cttttctgcc	420
ttgtggggag cctgaggact	ctggattctt	tgaggggatc ctggatgtgt	480
taaagagggct gttatcaggc	ttaaccataa	ccctcaagat ctgcttgaca	540
ccttagctca catccattcc	catcttcgg	gctccttagg cccaggatg	600
ggtccctgca agggccttt	cttgcacc	gcattca acca ttgataacca	660
tccctcttaa ggttcctct	acaacccaa	ggactttcat gattatcctc	720
ttggagggat tgagcgtgtt	tattaacaaa	ttgttttgg taataaaaata	780
aaaaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa aaaaaaaaaa	827

<210> 152
<211> 835
<212> DNA
<213> Homo sapiens

<400> 152

aaaatatttt ggtagtaatt	taaaatacaa	gaacgaatat	ttatttgc	acagttggag	60
atgttggata aatgtctttt	ctcaaagatc	acaggacttt	tgtcttcat	ttttgcctt	120
ttatattacca tttataaaag	atctggctcg	gattatggaa	tttaatgttt	atcagctcta	180
tgtattccctt tataaggct	tgaggaagta	tttcacataa	catgttttat	aatacttaac	240
catttatcca aagatataatt	tacattgggt	tgtccccctt	tcccttagat	catggtaaat	300
ttttcttatt gaggttaatta	tgtactactt	atatttqaag	gaagctttag	acattttaca	360
gtagctaaaa	tgttgagatt	agaggtactt	ttactattct	tctcaaagg	420
ataattaccc	aaattattca	agaaaataga	tcagaaataa	agaacaacat	480
gaattcattg	aaatttatgg	aatcagctct	cgcactgccc	atcttgcag	540
gaaattgctt	aatcacaaat	gttctacagt	ctttaatgt	agttagattt	600
tcatctgagt	aaattgattt	gtgattccag	agataagact	aatatttaa	660
atactgatta	gtataaaaaac	gtactcatca	cagaatttga	agcaaaatac	720
caaagagtaa	atgacaaatg	tataaatgct	gtagctcagg	attatatgt	780
tacactaata	aagattattt	ttcaaaaattt	aaaaaaa	aaaaaggcg	835

<210> 153
<211> 558
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (27)
<223> n equals a,t,g, or c

<220>
<221> SITE

<222> (39)

<223> n equals a,t,g, or c

<400> 153

cgggaccgga	taacaaattt	cacccnnga	aacaggctnt	gccccactag	gttttgca	60
aaaaagctat	tttaggttgc	cacttagga	gttacgcctg	gcaggtaccg	ggtccggaaa	120
ttcgcggccg	cgtccgactc	atgactgtgt	tggcaactta	aaaatattga	tatcccacaa	180
taaacaggg	tatcattgtat	ataatttccc	acatatttta	ctataaataa	tcgagtaaca	240
acctgtctt	taccattctt	tacagaaagg	ctttctcaa	tgcgttagtc	agggtttctt	300
cccggggaga	aaatttataa	tccttaatga	ggccagtact	cagaaggaca	tttctgctta	360
ctctttctc	tgtaattgcc	ctcaactaaaa	taaagcatga	ctttttatc	atgtgttac	420
acatgcagt	catccctaga	gttttctga	agcatgaatt	caataacata	taatttagacc	480
tgattctgag	aagattttct	cttcttcgtc	gacgcggccg	cgaatcccg	gtcgacgac	540
tcactagtcg	gcggccgc					558

<210> 154

<211> 1201

<212> DNA

<213> Homo sapiens

<400> 154

ggacatttgt	aaccctataa	acactagtaa	ataaaaaaca	gaaggacett	tatgtcctaa	60
catatctgtg	ttgtgaaagg	ctgcccgtg	aaatacggga	tttcttaaac	atattttaaa	120
aatcataggt	gtcaatattt	tttagaaatc	catttaaatt	ttctcttgc	attttacaat	180
gcctatttat	ttatatagtg	gctctgtga	tttgatgt	tatcctaaag	tttatatttt	240
ctttaaagga	tgtttatac	aactttatgt	aaaatgtttc	agtatctca	cattctctcc	300
ctgtcccttt	gttttgcct	tatatgtgg	tctgagtctt	ttctctggct	ttcaaaccta	360
gtaagactaa	gacactaaag	taacttgcc	cgagggttgg	gtaatgcck	cyaaakcaca	420
tcctaagctc	tcgtgcatac	aggggcctcc	tttgagctct	gtgctttga	gatcccatac	480
acctaaattc	cagtaactcca	aatcagtaact	gctcagtttt	agtgactaag	tttaaaaatg	540
tattnataa	rcaagttagt	ttagtgcctt	cttgcttctt	tctgactgc	ttgtatacat	600
gtatattcct	ttaaatgaat	cttggaaattt	atttagaaat	attaaattat	actaatgaaa	660
ctgtatattg	ttkgkaattc	ataagtgaaat	ttggaaagaa	tttgccttta	tgaaactaaa	720
tcctttttat	tcaagaatca	tatgtgtctt	tatatttatt	ccagcttaca	tttatatcac	780
tgagtaataa	tatagaaatg	tggatacata	cagctgttagt	tacagataca	aatatagata	840
taacctgtta	aatctatatc	tatcccataat	aacatatata	catgtaatat	gtgtgtgttt	900
atatatataat	gtttatgtca	ttaaagagct	cccttaatat	ttttctttta	tttcccttat	960
aatttgaggt	tgagcttga	tttccctgt	ataaaacaagg	aaatatttat	actagttta	1020
atactgtatgt	tttagacattt	tatcttattt	tagcgctgaa	tatttcaca	attattataa	1080
atattatcta	atactaataa	tgtacctgtt	aaaaatattt	aaaattttac	ctttgaatta	1140
ttttattgtt	gaattaaaat	tcctttaata	tgataaaaaaa	aaaaaaaaaa	aaaaactcgt	1200
a						1201

<210> 155

<211> 1026

<212> DNA

<213> Homo sapiens

<400> 155

gtctaaatgt	tcagtttcc	tccctaattc	caatgattct	cctcatttct	caatgtcctt	60
tgtccatctt	tgctgctcca	tttgcactgc	ctcccaaagg	tcactgtggc	tccttctctg	120
acttccacag	tcaagttaca	tttcataaaaa	attctaagct	cattttcaga	agccacaaat	180
ctatccttct	ttaaagtctt	caaacttga	ttgtgtaaat	aaatactcag	aaacaagatt	240
tctaaaaaaac	aaacactatt	ggccatcgta	tgtcaaaagg	agataacaaa	tgtttaacct	300
tatatgttgt	aggctttcta	aacttaattt	aaaaaaaaga	ctaaataaac	agtgtcaata	360
tgtctataaa	ctcacaacga	aaattttcag	atcatccat	tgtgtattca	ttggccggaa	420

acaatcatgt	aaaaaccaca	gcctggagc	tggtagcat	agaaaacaaga	480
tttcatggtt	ggtaactcaa	atctctaaag	ggktgtcagg	ttaaaaaaaaaa	540
gaaaagaata	gaaatttgac	ctgatctata	aaaatgaaag	tcgctggca	600
tttcactcc	tgacaaagat	gagctctc	ataggttagac	caaggcacac	660
ttcgtggcc	ccaaaattct	tcaagaaaat	agtagattga	ggaagcgatc	720
tagaggtgct	gttgaactg	gatgacattt	aagcttcctt	tttctccaa	780
ggccatgaag	catgctattt	catccccact	ccaaattgctg	tctccctggc	840
taccaccta	atcttgggtc	actgatctct	tttgcagaaa	atcagtccctg	900
gcaacttcat	cttccctaaaa	tgtcaccc	cttaaggcct	gctctgttca	960
cccagccaca	ccaatgtaaa	ctcggtccga	attcgatatac	aagtttatcg	1020
cctcgta				ataccgtcga	1026

<210> 156

<211> 904

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (8)

<223> n equals a,t,g, or c

<400> 156

acccacgn	ccggagtata	cttaatttta	tttatgtata	gaatatttgt	60
tggacatata	tttacactt	tgtcatttt	tttaaccaat	ttgagaatg	120
aattaattt	ttgcccggc	cttcataattt	tcttcttgc	tgcctctcc	180
gtactgttct	cacaatgcct	tttaaaaaatg	ttccataactg	tattagcata	240
acagaactaa	gaaatacatt	gctcaaataa	tattttactt	tattgataat	300
attttttaaa	ccccatcaaa	atagattca	attgactgtt	tcccctacat	360
acagtcgccc	atcgaataag	caaattgtt	tttgcagaata	aactgtaac	420
tgactctcag	aaggctttt	gctggatac	agaagagttt	ctaaaggctt	480
ttaataat	gttggtgagc	cagaggctt	acagagctgt	tactttagtg	540
attctcaggg	agttagttat	tcatcattt	gtaaagcccct	ccccacactc	600
aacaagtat	gaaggcttat	cttaaaactgt	gtagtagctt	agacttggca	660
atagagcaga	gataaaat	tttgcattt	gaaatcaat	tttctgttac	720
aaaattttat	tttctggaa	attatata	ttcaaaatgt	tgatgttg	780
ttggttacaa	gagaataatg	ttacatgtt	aattgttata	tttgcctt	840
tccctttcag	tcataataaa	tgatttacaa	aacccaaaaa	aaaaaaa	900
ggcc				aaaaaggc	904

<210> 157

<211> 916

<212> DNA

<213> Homo sapiens

<400> 157

gtttgtttaa	ccatgttctt	cagaatgcag	gtatgtgagc	atcatggttt	60
ctgctgtcc	tgtctttgaa	aatggagata	ccacttgcag	cttacccac	120
tccagcat	gtatgtgttt	cactccattt	catccatcca	gaactttcac	180
ccattacca	gcattttta	acattgatca	ataaggccta	taaccagatt	240
acaccagagg	tctggggca	agggtggaaa	ttgactttac	attcttagta	300
cataagtgt	ttatatatat	attgttgtt	ttgatcatct	attcaaaaaa	360
gcagctgt	ttgtataggc	tctgtgtt	ccagtgaaga	tacatgatta	420
tcagagctgt	gcacagcaac	acacagaagg	agagataact	cacccagctt	480
ggacagagaa	tgaggttagc	ctcccagat	tccttgc	agttttagct	540
				gttttcaggt	600

gttgataaaaa gctccaggag ctggcaggag gagagcagag gaagcctcgag cttacaaagc	660
acaaaaggcca tgacagcatg ccagacgggt gaaagaggac agggaaaatg taggcaagtg	720
tctcttctca gaggatgtta tatactatgt taaaaagtgt tgatctgctg ggcacagtgg	780
ttcacgcattg tagtgtcagc actttgggtt gccaagggtgg gaggattgtctgagtcagg	840
agtttgagac cagcctggc aacatagtga gaccccatct cttaaaaaaaaaaaaaaaaaaaa	900
aaaaaaaaaggc gcgcgc	916

<210> 158
<211> 921
<212> DNA
<213> Homo sapiens

<400> 158	
ggaactgctg ctcatggAAC tggctccTCT cctttGCCA cttgagtctg ttgcagaAGT	60
ccaggGAAGA acTTGAAGAG caAAATACAC tcttGAGTT gttGGGTTT gggAGAGGTG	120
acAGTAGAGA agggGGTTGt gtTTAAAATA AACACAGTGG cttgAGCAGG ggcAGAGGTT	180
gtgatGCTAT ttctGTTGAC tcctAGCAGC catcACCAGC atGAATGTGT tcgtAGGGCC	240
tttGAGTGTG GCGATTGTCA tattctGTTG gataACAATG tattGGGTGT cgattGTcat	300
ggggcAGGGG agAGGGCAGT acACCTGGAG gaccATTTG tccACATCGA caccATCAGT	360
ctgctCTTAg aggATGCCCt ggAGTATTGc gcGTTGATTG cggggCACCC gaaATCAGAC	420
ttGCCACCTG gACTGTCGAG gtGcAGACCC tggGAGCACC ACTGGCCAT ctcttACACa	480
ggctGACCgA tttctCCTGG tGTTcAGAGT ctGTTTTGT ctAGCACCt ttGAAATCgg	540
ttatGATGTA gggggAAAAG cAGCAGCCTC gaAGCCTCAT gccaACTCTG ggcAGCAGCA	600
gcctGTTGGT tcctGGAAGA tggatGGGCA gagaATAGGG aAGGAAGATC atGCTTTCC	660
ctactAACTT ctGTAActGC atGTATGATA cATTATTGCA gagGTAAGAG atAGTTAAT	720
gGATTTTAA aaACAAATTa ctATAATTa tctGATGTTc tctAGTTGCA ttttGCTGAA	780
atGTagtGtC gttctAAATT CTGTAATTG attGCTGTG aattatCTTt ctGTTGAGAA	840
aaaaaaaAGGGCGCG	900
aaaaaaaaaaaa aGGGCGGCGC	921

<210> 159
<211> 804
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (800)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (801)
<223> n equals a,t,g, or c

<400> 159	
aagaaaaACTC tagaAGTCTG gATGTCGGTG ggcCTCTGAG ATATGCCGTT tacGGTTCTT	60
cttcACAGGG CGCTGAGTC acTTCTTCTA ctttCTTCTG GAACATTGGA tccCTCCTGA	120
ggTCCCCCTG GCAGGGCTCA ggAGGCTTCT CTCGGACCGC CTCGTTGTTG cacGGCCTT	180
cctCATGTTG ttcttCCTCA tCATGAACCT TCTGGAGGGG AAAGACGCC CAGCCTTCGC	240
CGCCAAGATG AGGGGGGGCT TCTGGCCGGC GCTGAGGATG AACTGGCGGG TGTGGACGCC	300
actACAGTTC ATCAACATCA ACTACGTCCC TCTGAAGTTC CGGGTCTCT TCGCCAACCT	360
GGCAGCTTG ttctGGTATG CCTACCTGGC CTCCTGGGG AAGTGCAGCAG CGCTGGGAGA	420
ACATCAGGTG CACTGTGGAC GTGGGTCTGG GGGTCTCACC CGCCCAAGCAG GAGCAGAAC	480
AATCCAGTCA GGATGTCACT GACTCTAAAT CAGGTGATTc AAGATGCCA AAAATGATGG	540
ATAGAGAAAC AGAAATCTCT GAATGTCAAGA ACCCTGTCTT TTTAAAAAGGC AGTCRCTGCC	600

ttcaggtgg tctgccccag aaacttaaaa tttagtcgag gcagttcaa ttttactgt 660
 ggaccgaatt aggatcacaa taaacgataa tgcaggttct tcaaaaaaaaaaaaaaa 720
 aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaactc gagggggggc ccgtacccaa 780
 tcgccctgat gatgatctgn ncac 804

<210> 160
 <211> 24
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (24)
 <223> Xaa equals stop translation

<400> 160
 Met Tyr Gly Cys Val Cys Val Cys Ile Tyr Leu Tyr Thr Cys Ile His
 1 5 10 15
 Gly Cys Pro Cys Val Ser Met Xaa
 20

<210> 161
 <211> 113
 <212> PRT
 <213> Homo sapiens

<400> 161
 Met Gly Ser Trp Cys Ile Cys Thr Leu Leu Leu Leu Thr Asp Gly
 1 5 10 15

Gln Gln Gly Phe Tyr Pro Gln Pro Phe Gln Ala Ala Pro Gly Arg Gln
 20 25 30

Gln Leu Trp Gly Gly Thr Asn Pro Trp Ala Val Leu Ile Pro Glu Ser
 35 40 45

Phe Leu Pro Tyr Thr Leu Thr Val Asn Tyr Ser Pro Ser Cys Asn Phe
 50 55 60

Glu Phe Tyr Leu Pro Lys Met Arg Leu Ala Tyr Ile Cys Met Ser His
 65 70 75 80

Ser His Cys Pro Tyr Leu Gly Arg Asp Ile Ile Ile Thr Leu Leu Asn
 85 90 95

Tyr Cys Ser Ser Phe Leu Ala Glu Leu Leu Ala His Leu Val Tyr Ile
 100 105 110

Ala

<210> 162
 <211> 45
 <212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (45)

<223> Xaa equals stop translation

<400> 162

Met	Thr	Lys	Arg	Arg	Lys	Pro	Arg	Tyr	Arg	Phe	Ile	Phe	Ala	Leu	Tyr
1															15

Ala	Leu	Arg	Leu	Val	Phe	Leu	Phe	Arg	Ala	Val	Thr	Asn	Thr	Asp	Ala
															30
	20					25									

Ser	Arg	Leu	Arg	Ala	Lys	Arg	Gly	Glu	Cys	Pro	Tyr	Xaa			
															45
	35					40									

<210> 163

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals stop translation

<400> 163

Met	Thr	Glu	Gly	Leu	Leu	Ser	Ser	Leu	Ser	Leu	Leu	Tyr	Leu	Leu
1														15

Thr	Trp	Leu	Leu	Met	Leu	Ser	Lys	Lys	Leu	Tyr	Val	Gln	Met	Ile	Phe
															30
	20					25									

Cys	Tyr	Asn	Pro	His	Phe	Ser	Gln	Met	Asp	Ala	Cys	Asn	Gly	Thr	Ser
															45
	35					40									

Gln	Lys	Ile	His	Asn	Ala	Arg	Gln	Cys	Thr	Xaa					
	50					55									

<210> 164

<211> 118

<212> PRT

<213> Homo sapiens

<400> 164

Met	Cys	Tyr	Leu	Leu	Leu	Leu	Ile	Gln	Thr	Ala	Glu	Leu	Leu	Ile
1														15

His	Pro	Gln	Gly	Leu	Gln	Ala	Val	Ser	Asn	Gly	Glu	Ser	Ala	Leu	Lys
															30
	20					25									

Gly	Thr	Arg	Pro	Thr	Phe	Ser	Ser	Pro	Phe	Ile	Leu	Val	Thr	Glu	Gly
															45
	35					40									

Arg Lys Glu Trp Glu Gly Val Phe Leu Ser Ser Gly Trp Lys Gly Asn

50

55

60

Thr Leu Ser Asn Tyr Tyr Ile Ser Leu Val Phe Tyr Tyr Ser Arg Ile
65 70 75 80

Leu Gln Pro Tyr Phe Tyr Cys Leu Trp Gly Lys Leu Glu Met Val Thr
85 90 95

Leu Ile Arg Ser Val Trp Arg Gly Ile Asn Gly Gly Asp Lys Ile Ser
100 105 110

Val Gly Phe Gly Lys Cys
115

<210> 165

<211> 55

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (55)

<223> Xaa equals stop translation

<400> 165

Met Cys Ser Gly Leu Leu Ser Met Thr Phe Ser Phe Leu Leu Glu Phe
1 5 10 15

Cys Ser Val Ala Gln Arg Leu Arg Leu Ala Asp Ala Arg Thr Ser Met
20 25 30

Gln Asp Ile Leu Lys Trp Phe Ser Asp Tyr Thr Leu Arg Ala Asp Ile
35 40 45

Ser Lys Ser Arg Asp Leu Xaa
50 55

<210> 166

<211> 127

<212> PRT

<213> Homo sapiens

<400> 166

Met Gln Gly Ser Asp Ala Gly His Gly Gly Thr His Ile Tyr Arg Ala
1 5 10 15

Leu Val Gln Trp Pro Leu Ala Trp Val Phe Tyr Leu Ser His Ala Lys
20 25 30

Thr His Trp Gly Glu Glu Leu Arg Phe Ser Phe Arg Arg Lys Asn Leu
35 40 45

Arg Leu Arg Glu Ala Met Arg His Glu Thr Cys Gln Val Thr Gln Leu
50 55 60

Val Ala Gly Lys Ala Asp Ser Asn Leu Cys Leu Arg Asp Ser Glu Thr

65 70 75 80

Trp Phe Trp Pro Pro Leu Trp Ala Ala Cys Ser Ser Leu Gln Ala Thr
85 90 95

Ala Cys Arg Leu Ser Ser Pro Ser Lys Gly Leu Gly Ala Ser Arg Glu
 100 105 110

Cys Pro Trp Leu Ala Ser Gly Arg Ala Ala Leu Val Ser Phe Leu
 115 120 125

<210> 167

<211> 56

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (32)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 167

Met	Gly	Val	Glu	Gln	Tyr	Ser	Tyr	Leu	Phe	Leu	Thr	Cys	Val	Phe	Met
1				5				10						15	

Cys Val Ser Leu Gln Trp Lys Ser Thr Gln Pro Trp Val Gly Asp Xaa
 20 25 30

Thr Cys Met Arg Lys Gly Ile Thr Gly Thr Glu Val His Arg Thr Asn
35 40 45

Ala Leu Phe Thr Phe Trp Cys Ser
50 55

<210> 168

<211> 73

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (73)

<223> Xaa equals stop translation

<400> 168

```

Met Pro Ser Ile Arg Leu Gly Leu Ser His Leu Phe Leu Thr Ala Gly
      1           5           10          15

```

Ile Tyr Cys Leu Leu Leu Cys Ala Arg Cys Cys Ala Leu Gly Arg Gly
20 25 30

Thr Ala Trp Ala Ala Cys Pro Gly Gly Ala Cys Gly Leu Met Gly Glu
35 40 45

Ala Asp Pro Ser Pro Pro His Cys Gln Gln Gly Gln Gly Lys Ser Thr
50 55 60

His Arg Gly Leu Ile Pro Tyr Val Xaa
65 70

<210> 169

<211> 70

<212> PRT

<213> Homo sapiens

<400> 169

Met Thr Pro Gln Asn Leu Arg Phe Thr Leu Phe Gln Phe Cys Tyr Ser
1 5 10 15

Leu Tyr Leu Glu Leu Glu Leu Gly Phe Arg Ser Leu Ser Gln Glu Val
20 25 30

Thr Arg Glu Trp Cys Leu Ser Tyr Phe Phe Leu Ile Lys Val Cys Trp
35 40 45

Gln Val Pro Val Ser Glu Phe Leu Leu Val Lys Glu Asn Pro Phe Leu
50 55 60

Leu Leu Glu Lys Lys Leu
65 70

<210> 170

<211> 80

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (80)

<223> Xaa equals stop translation

<400> 170

Met Pro Phe Ile Leu Leu Leu Val Cys Leu Thr Ser Leu Pro Ser Arg
1 5 10 15

Gly Tyr Asn Glu Lys Lys Leu Thr Asp Asn Ile Gln Cys Glu Ile Phe
20 25 30

Gln Val Leu Tyr Glu Glu Ala Thr Ala Ser Tyr Lys Glu Glu Ile Val
35 40 45

His Gln Leu Pro Ser Asn Lys Pro Glu Glu Leu Glu Asn Asn Val Asp
50 55 60

Gln Ile Leu Lys Trp Ile Glu Gln Trp Ile Lys Asp His Asn Ser Xaa
65 70 75 80

<210> 171

<211> 42
<212> PRT

<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 171

Met Lys Ile Leu Ile Leu Phe Ile Phe Ile Pro Gly Leu Leu Val Glu
1 5 10 15

Lys Asn Gly Pro Asp His Val Cys Val Cys Met Cys Val Arg Val Cys
20 25 30

Val Cys Ala His Leu Gly Leu Phe Ile Xaa
35 40

<210> 172
<211> 131
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (43)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (44)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (49)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (66)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (78)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (94)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (102)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 172

Met Trp Ser Val Ile Arg Ser Leu Cys Pro Ser Arg Leu Gln Ser Leu
1 5 10 15

His Val Cys Phe Cys Pro Arg Leu Cys Leu Ala Val Pro Cys Val Phe
20 25 30

His Leu Ser Ser Pro Trp Phe His Val Arg Xaa Xaa Phe Phe Ser Gly
35 40 45

Xaa Pro Gly Cys Ile Trp Gly Ile Cys Phe Val Gly Leu Leu Gly
50 55 60

Ala Xaa Arg Pro Arg Ser Gly Cys Leu Cys Ser Pro Ser Xaa Cys Leu
65 70 75 80

Trp Ser Leu Val Val Cys Glu Ser Ile Cys Leu Pro Arg Xaa Gly Pro
85 90 95

Asn Gln Ala Pro Pro Xaa Pro Leu Phe Leu Ser Leu Asn Leu Pro Phe
100 105 110

Leu Phe Gln Pro Leu Gln Met Arg Trp Leu Ser Ala Val Gly Trp Arg
115 120 125

Glu Ala Met
130

<210> 173

<211> 45

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (45)

<223> Xaa equals stop translation

<400> 173

Met Gln Leu Ser Leu Ser Leu Cys Ala Phe Val Val Cys Thr Asn Ala
1 5 10 15

Val Cys Thr His Ala Ala Thr Asn Gln Ala Arg Leu Val Gly Phe Leu
20 25 30

Lys Val Leu Arg Pro Ala His Ser Pro Leu Cys Leu Xaa
35 40 45

<210> 174

<211> 63

<212> PRT

<213> Homo sapiens

<220>

<221> SITE
<222> (10)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (25)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (38)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (63)
<223> Xaa equals stop translation

<400> 174
Met Gln Pro Ala Trp Leu Trp Leu Trp Xaa Trp Glu Leu Gly Trp Glu
1 5 10 15

Leu Val Phe Gly Ala Ile Leu Leu Xaa Leu Gln Asp Gly Leu Phe Asp
20 25 30

Ser Val Leu Tyr Cys Xaa His Leu Tyr Ser Gly Leu Phe Phe Pro Trp
35 40 45

Ile Val Asn Ser Leu Met Ser Gly Ser Ser Gln Leu Met Ser Xaa
50 55 60

<210> 175
<211> 20
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (20)
<223> Xaa equals stop translation

<400> 175
Met Ser Ser Pro Lys Ser Leu Val Leu Leu Leu Ala Val Ile Cys Arg
1 5 10 15

Glu Ala Thr Xaa
20

<210> 176
<211> 41
<212> PRT
<213> Homo sapiens

<220>
<221> SITE

<222> (41)

<223> Xaa equals stop translation

<400> 176

Met Asn Ile Val Pro Gln Phe Ser Val Leu Pro His Phe Ala Tyr Phe
1 5 10 15

Ser Phe Ile Ile Leu Tyr Trp Ala Val Leu Phe Ser Gln Thr Ile Cys
20 25 30

Ser Met Ser Val Phe Lys Val Lys Xaa
35 40

<210> 177

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals stop translation

<400> 177

Met Thr Asp Ile Thr Cys Phe Leu Phe Ser Tyr Leu Ser Thr Leu Leu
1 5 10 15

Ser Pro Ile Tyr Leu Asp Val Leu Leu Phe Ser Leu Leu Leu Phe Leu
20 25 30

Phe His Ile Ala Gly Met His Ile Leu Thr Phe Ile Asn His Asp Ile
35 40 45

Xaa

<210> 178

<211> 107

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (63)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (65)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (77)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (88)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (105)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (107)

<223> Xaa equals stop translation

<400> 178

Met	Gly	Ala	Ala	Leu	Ala	Ala	Trp	Ile	Cys	Ile	Val	Arg	Tyr	His	Gln
1				5						10					15

Leu	Arg	Asp	Trp	Gly	Val	Arg	Arg	Trp	Pro	Asn	Gln	Leu	Ile	Leu	Trp
				20				25						30	

Thr	Gly	Leu	Leu	Cys	Ala	Leu	Gly	Thr	Ser	Val	Val	Gly	Asn	Leu	Pro
				35				40					45		

Gly	Glu	Thr	Gln	Ser	Ala	Pro	Arg	Val	Cys	Xaa	Arg	Pro	Ala	Xaa	Gly
	50					55					60				

Xaa	Thr	Thr	Pro	Ser	Met	Pro	Arg	Gly	His	Arg	Leu	Xaa	Val	Ser	Gly
	65				70				75				80		

Ala	Gly	Ser	Arg	Pro	Pro	Phe	Xaa	Gly	Leu	Val	Phe	Phe	Ser	Gly	His
					85				90				95		

Trp	Pro	Gly	Pro	Ala	Gly	Ser	Phe	Xaa	Leu	Xaa					
				100				105							

<210> 179

<211> 46

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (46)

<223> Xaa equals stop translation

<400> 179

Met	Gly	Cys	Trp	Val	Leu	Phe	Ile	Leu	Leu	Tyr	Leu	Ala	Leu	His	Ile
1				5					10					15	

Cys Val Gln Asn Tyr Ile Tyr Ser Tyr Lys Ile Ile Cys Leu Gln Ser

20

25

30

Phe His Tyr Ile Val Arg Lys Ile Gln Ile Phe Val Ser Xaa
35 40 45

<210> 180
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (67)
<223> Xaa equals stop translation

<400> 180
Met Leu Leu Ala Ala Phe Leu Ala Leu Phe Pro Leu His Asp Ser Arg
1 5 10 15

Gly Leu Lys His Thr Gly Ala Gly His Val Asn Ser Val Ala Leu Leu
20 25 30

Pro Ile Pro Leu Lys Ala Val Ser Leu Ser Pro Val Ser Ser Leu Gln
35 40 45

Val Pro Cys Cys Cys Ser Ser Phe Gln Leu Leu Leu Thr Phe Leu Ser
50 55 60

Val Ser Xaa
65

<210> 181
<211> 50
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation

<400> 181
Met Ile Cys Lys Phe Leu Ile Ile Cys Ile Thr Leu Leu Leu Phe
1 5 10 15

Ala Ile Cys Gln Leu Cys Lys Arg Gln Gly Leu Val Gln Lys Ile Ser
20 25 30

Phe Tyr Gln Lys Glu Thr Leu Ser Ser Thr Val Gly Thr Thr Phe Leu
35 40 45

Ser Xaa
50

<210> 182

<211> 73
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (73)
<223> Xaa equals stop translation

<400> 182
Met Leu Thr Trp Val Trp Tyr Leu Ile Met Thr Ser Val Leu Gln Ala
1 5 10 15

Ser Val Ser Ser Val Val Arg Gly Ser Ile Leu Val Gly Gly Ser Glu
20 25 30

Asp Cys Xaa Glu Gly Gly Ser Leu Ile Gln Val Ser Leu Gly Tyr Val
35 40 45

Leu Ala Ala Arg Glu Asp Arg Gln Glu Cys Gly Pro Asp Thr Val Ser
50 55 60

Cys Pro Pro Gly Met Arg Leu Asp Xaa
65 70

<210> 183
<211> 44
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation

<400> 183
Met Leu Ser Ala Leu Ser Ala Leu Tyr Leu Ile Ile Thr Ile Phe Leu
1 5 10 15

Lys Gly Ser Cys Cys Ser Cys His His Cys Phe Thr Asn Gly Lys Leu
20 25 30

Trp Leu Arg Lys Phe Ile Ser Gly Ser Gln Pro Xaa
35 40

<210> 184
<211> 58
<212> PRT
<213> Homo sapiens

<220>

<221> SITE
 <222> (58)
 <223> Xaa equals stop translation

<400> 184
 Met Cys Met Thr Val Phe Ile Val Phe Tyr Tyr Ser Phe Met Arg Leu
 1 5 10 15

Leu Phe Arg Cys Ser His Asn Arg Arg His Trp Arg Gly Ser Gly Lys
 20 25 30

Asn Thr Val Tyr His Thr Gly Pro Arg Asp Glu Ala Cys Cys Ala Met
 35 40 45

Pro Cys Trp Ala Thr Trp Gly Arg Arg Xaa
 50 55

<210> 185

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (69)

<223> Xaa equals stop translation

<400> 185

Met Pro Leu Ala Leu Lys Arg Gly Gln Leu Phe Leu Ile Pro Trp Leu
 1 5 10 15

Phe Pro Gln Gly Val Cys Pro Leu Glu Gly Glu Gln Leu Gly Ser Gly
 20 25 30

Lys Glu Gly Leu Leu Gln Phe Ala Ile Ala Ser Cys Pro Arg Val Tyr
 35 40 45

Pro Glu His Ser Pro Pro Trp Lys Glu Thr Gln Ser Ala Thr Gly Tyr
 50 55 60

Arg Lys Ser Asp Xaa
 65

<210> 186

<211> 25

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (25)

<223> Xaa equals stop translation

<400> 186

Met Lys Tyr Leu Leu Phe Leu Val Phe Cys Leu Ser Tyr Val Lys Asp
 1 5 10 15

Leu Asn Ile Phe Asp Leu Leu Tyr Xaa
20 25

<210> 187
<211> 58
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (58)
<223> Xaa equals stop translation

<400> 187
Met Thr Leu Pro Trp Glu Trp Val Pro Asp Lys Arg Ile Trp Leu Leu
1 5 10 15

Ser Leu Thr Leu Val His Ala Leu Leu Pro Leu Cys Leu Leu Pro Trp
20 25 30

Asp Val Gly Ala Arg Ser Pro Phe Ile Ser Gly Glu Pro Ile Asn Leu
35 40 45

Gly Phe Pro Asn Leu Gln Asn Cys Lys Xaa
50 55

<210> 188
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (67)
<223> Xaa equals stop translation

<400> 188
Met Val Gly Leu Leu Leu Ile Ala Leu Leu Thr Trp Gly Tyr Ile Arg
1 5 10 15

Tyr Ser Gly Gln Tyr Arg Glu Leu Gly Gly Ala Ile Asp Phe Gly Ala
20 25 30

Ala Tyr Val Leu Glu Gln Ala Ser Ser His Ile Gly Asn Ser Thr Gln
35 40 45

Ala Thr Val Arg Asp Ala Val Val Gly Arg Pro Ser Met Asp Lys Lys
50 55 60

Ala Gln Xaa
65

<210> 189
<211> 89

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (18)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (63)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (89)

<223> Xaa equals stop translation

<400> 189

Met Ser Thr Tyr Leu Lys Met Phe Ala Ala Ser Leu Leu Ala Met Cys
1 5 10 15

Ala Xaa Ala Glu Val Val His Arg Tyr Tyr Arg Pro Asp Leu Met Arg
20 25 30

Asn Arg Leu Arg Arg Val Lys Leu Ile Ser Gln Ser His Ile Ala Leu
35 40 45

Val Arg Arg Phe Glu Asp Leu Lys Pro Lys Leu Ser Val Cys Xaa Thr
50 55 60

Gly Ile Thr Ser Leu Ser Val Gly Glu Leu Glu Val Trp Ala Glu Ser
65 70 75 80

Ser Arg Gly Asp Leu Met Thr Ala Xaa
85

<210> 190

<211> 221

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (159)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (221)

<223> Xaa equals stop translation

<400> 190

Met Lys Leu Leu Leu Trp Ala Cys Ile Val Cys Val Ala Phe Ala Arg
1 5 10 15

Lys Arg Arg Phe Pro Phe Ile Gly Glu Asp Asp Asn Asp Gly His

20

25

30

Pro Leu His Pro Ser Leu Asn Ile Pro Tyr Gly Ile Arg Asn Leu Pro
35 40 45

Pro Pro Leu Tyr Tyr Arg Pro Val Asn Thr Val Pro Ser Tyr Pro Gly
50 55 60

Asn Thr Tyr Thr Asp Thr Gly Leu Pro Ser Tyr Pro Trp Ile Leu Thr
65 70 75 80

Ser Pro Gly Phe Pro Tyr Val Tyr His Ile Arg Gly Phe Pro Leu Ala
85 90 95

Thr Gln Leu Asn Val Pro Pro Leu Pro Pro Arg Gly Phe Pro Phe Val
100 105 110

Pro Pro Ser Arg Phe Phe Ser Ala Ala Ala Ala Pro Ala Ala Pro Pro
115 120 125

Ile Ala Ala Glu Pro Ala Ala Ala Ala Pro Leu Thr Ala Thr Pro Val
130 135 140

Ala Ala Glu Pro Ala Ala Arg Gly Pro Val Ala Ala Glu Pro Xaa Gly
145 150 155 160

Arg Gly His Leu Leu Glu Leu Glu Pro Ala Ala Glu Ala Pro Val Ala
165 170 175

Ala Glu Pro Ala Ala Glu Ala Pro Val Gly Val Glu Pro Ala Ala Glu
180 185 190

Glu Pro Ser Pro Ala Glu Pro Ala Thr Ala Lys Pro Ala Ala Pro Glu
195 200 205

Pro His Pro Ser Pro Ser Leu Glu Gln Ala Asn Gln Xaa
210 215 220

<210> 191

<211> 52

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (52)

<223> Xaa equals stop translation

<400> 191

Met Glu Arg Leu Val Leu Ser Leu Trp Ser Leu Thr Cys Arg Ala Ser
1 5 10 15

Pro Ala Asn Thr His Pro Arg Thr Thr Ser Arg Thr Arg Thr Leu Asp
20 25 30

Val Lys Thr Lys Cys Pro Val Glu Ala Val Lys Leu Ser Glu Met Leu
35 40 45

Pro Pro Val Xaa
50

<210> 192
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (72)
<223> Xaa equals stop translation

<400> 192
Met Val Gly Thr His Leu Ile Leu Phe Pro Phe Leu Leu Arg Thr Met
1 5 10 15

Val Ile Phe Leu Cys Leu Lys Ser Ser Cys Gly Ser Phe Leu Pro Ile
20 25 30

Asn Lys Ile Gln Thr Pro Phe Ile Leu Asn Leu Ile Tyr Lys Thr Phe
35 40 45

Lys Met Cys Ser Leu Pro Asn Ser Leu Phe Ser Pro Leu Ser Phe Ile
50 55 60

Phe Phe Ile Phe Phe Leu Thr Xaa
65 70

<210> 193
<211> 112
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (108)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (112)
<223> Xaa equals stop translation

<400> 193
Met Arg Arg Leu Leu Leu Ala Leu Pro Phe Ala Leu Leu Pro Leu Ala
1 5 10 15

Val Ala His Ala His Glu Asp His Asp His Glu His Gly Ser Leu Gly
20 25 30

Ala His Glu His Gly Val Gly Arg Leu Asn Ala Val Leu Asp Gly Gln
35 40 45

Ala Leu Glu Leu Glu Leu Asp Ser Pro Ala Met Asn Leu Val Gly Phe

50

55

60

Glu His Val Ala Thr Ser Ala Ala Asp Lys Ala Lys Val Ala Ala Val
65 70 75 80

Arg Lys Gln Leu Glu Asn Pro Ser Gly Pro Val Gln Pro Ala Gln Ser
85 90 95

Arg Ser Cys Val Val Ser Asn Gln Gly Ile Asn Xaa Arg Cys Ser Xaa
100 105 110

<210> 194

<211> 61

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (14)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (61)

<223> Xaa equals stop translation

<400> 194

Met Phe Ile Thr Arg Gly Cys Tyr Cys Phe Val Phe Phe Xaa Leu Ala
1 5 10 15

His Asn Cys Lys Ala Ala Arg Thr Thr Arg Asn Gly Phe Pro Thr Val
20 25 30

Pro Gly Arg Arg Gln Arg Thr Leu Arg Arg Leu Phe Leu Cys Gly Phe
35 40 45

Pro Leu Leu Cys Ser Gln Gly Asp Leu Ser Ala Ala Xaa
50 55 60

<210> 195

<211> 126

<212> PRT

<213> Homo sapiens

<400> 195

Met Thr Lys Leu Ala Gln Trp Leu Trp Gly Leu Ala Ile Leu Gly Ser
1 5 10 15

Thr Trp Val Ala Leu Thr Thr Gly Ala Leu Gly Leu Glu Leu Pro Leu
20 25 30

Ser Cys Gln Glu Val Leu Trp Pro Leu Pro Ala Tyr Leu Leu Val Ser
35 40 45

Ala Gly Cys Tyr Ala Leu Gly Thr Val Gly Tyr Arg Val Ala Thr Phe
50 55 60

His Asp Cys Glu Asp Ala Ala Arg Glu Leu Gln Ser Gln Ile Gln Glu
65 70 75 80

Ala Arg Ala Asp Leu Ala Arg Arg Gly Cys Ala Ser Asp Ser Leu Thr
85 90 95

Pro Phe Leu Cys Gly Gln Pro Phe Leu Pro Phe Pro Ile Lys Glu Pro
100 105 110

Val Tyr Phe Leu Lys
115 120 125

<210> 196

<211> 113

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (41)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (109)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (113)

<223> Xaa equals stop translation

<400> 196

Met Ala Ala Leu Leu Leu Pro Trp Leu Met Leu Leu Thr Gly Arg
1 5 10 15

Val Ser Leu Ala Gln Phe Ala Leu Ala Phe Val Thr Asp Thr Cys Val
20 25 30

Ala Gly Ala Leu Leu Cys Gly Ala Xaa Leu Leu Phe His Gly Met Leu
35 40 45

Leu Leu Arg Gly Gln Thr Thr Trp Glu Trp Ala Arg Gly Gln His Ser
50 55 60

Tyr Asp Leu Gly Pro Cys His Asn Leu Gln Ala Ala Leu Gly Pro Arg
65 70 75 80

Trp Ala Leu Val Trp Leu Trp Pro Phe Leu Ala Ser Pro Leu Pro Gly
85 90 95

Asp Gly Ile Thr Phe Gln Thr Thr Ala Asp Val Gly Xaa Thr Ala Ser
100 105 110

Xaa

<210> 197
<211> 66
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (66)
<223> Xaa equals stop translation

<400> 197
Met Leu Gly Ile Thr Arg Leu Trp Val Leu Leu Lys Pro Cys Phe Pro
1 5 10 15

Arg Cys Tyr Ser Ser Thr Gly Gly Glu Val Leu Pro Arg Cys Cys Glu
20 25 30

Val Glu Ala Glu Val Gln Val Pro His Ser Ala Pro Met Asp Ser Arg
35 40 45

Glu Gly Gly Thr Val Pro Tyr Phe Gly Gly Cys Ser Pro Arg Phe
50 55 60

Tyr Xaa
65

<210> 198
<211> 52
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (23)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (52)
<223> Xaa equals stop translation

<400> 198
Met Ala Gln His His Leu Leu Ser Ile Leu Leu Ala Ile Leu Ser Cys
1 5 10 15

Ser Ser Gln Pro Arg Gln Xaa Arg Gly Ser Gly Ala Leu Pro Cys Glu
20 25 30

Val Cys Ser Ala Val Leu Leu Thr Cys Leu Arg Lys Ile Ser Gly Ser
35 40 45

Leu Cys Val Xaa

50

<210> 199
<211> 59
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (59)
<223> Xaa equals stop translation

<400> 199
Met Ile Gly Lys Ser Leu Val Met Phe Cys Phe Leu Ser Trp Gly Ala
1 5 10 15

Gly Val His Gly Cys Ala Leu Tyr Tyr Asn Ala Ser Asn Arg Ile Gly
20 25 30

Ile Phe Tyr Ile Phe Cys Phe Thr Tyr Leu Arg Leu His Glu Cys Val
35 40 45

Met Leu Ser Asn Leu Arg Val Asn Glu Leu Xaa
50 55

<210> 200
<211> 52
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (52)
<223> Xaa equals stop translation

<400> 200
Met Leu Ser Pro Leu Ser Gln Ser Leu Leu Val Ala Leu Asn Val Leu
1 5 10 15

Phe Leu Leu Pro Asn Phe Leu Ala Leu Ser Lys Asn Leu Thr Tyr Asp
20 25 30

Cys Tyr Phe Arg Phe Phe Pro Thr Phe Phe Leu Pro Pro Lys Glu Met
35 40 45

Trp Tyr Leu Xaa
50

<210> 201
<211> 81
<212> PRT
<213> Homo sapiens

<220>
<221> SITE

<222> (81)

<223> Xaa equals stop translation

<400> 201

Met Cys Pro Ala Ala Ala Leu Ala Trp Pro Thr Ser Ala Ile Ser Leu
1 5 10 15

Ile Val Ser Leu Ala Pro Ser Trp Ala Ala Ala Arg Asp Asn Trp Ala
20 25 30

Ala Ser Pro Tyr Thr Thr Gln Ala Arg Pro Ala Leu Arg Ala Ala Leu
35 40 45

Thr Thr Ile Ser Gly Pro Met Pro Ala Ala Ser Pro Met Val Met Pro
50 55 60

Thr Gly Arg Glu Gly Phe Thr Val Leu Gly Met Gly Leu Arg Cys Gly
65 70 75 80

Xaa

<210> 202

<211> 70

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (70)

<223> Xaa equals stop translation

<400> 202

Met Phe Leu Ile Val Phe Cys Phe Leu Gln Ser Leu Ser Ala Met Pro
1 5 10 15

Ile Val Leu Ile Phe Tyr Arg Ser Ser Leu Lys Ile Leu Asn Arg Gly
20 25 30

Ile Gly Ser Gly Gln Ser Glu Trp Leu Glu Phe Trp Leu Ser Lys Lys
35 40 45

Asn Phe Ile Leu His Lys His Val Val Arg Ser Phe Cys Ala Tyr Ala
50 55 60

Ala Trp Ile Gly Cys Xaa
65 70

<210> 203

<211> 46

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (46)

<223> Xaa equals stop translation

<400> 203

Met Leu Leu Cys Ser Val Arg Asn Ile Leu Trp His Thr Ala Phe Leu
1 5 10 15

Gly Ser Ala Val Leu Cys Phe Val Leu Val Leu Val His Leu Glu
20 25 30

Cys Leu Ile Ile Asp Ala Tyr Phe Asn Ser Ile Ser Phe Xaa
35 40 45

<210> 204

<211> 53

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (53)

<223> Xaa equals stop translation

<400> 204

Met Gly Thr Glu Ala Ser Pro Lys Arg Tyr Phe Phe Val Val Val
1 5 10 15

Val Leu Gly Ile Ile Val Pro Ile Leu Arg Ala Phe Pro Pro Pro Val
20 25 30

Pro Thr His Pro Asn Lys Met Trp Trp Cys Cys Leu Gln Lys Arg Glu
35 40 45

Val Leu Cys His Xaa

50

<210> 205

<211> 62

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (62)

<223> Xaa equals stop translation

<400> 205

Met Phe Cys Trp Ile Leu Val Cys Leu Ala Tyr Leu Lys Val Pro Leu
1 5 10 15

Leu Phe Phe Phe Phe Phe Leu Ser Ala Leu Phe Cys Arg Thr Cys
20 25 30

Ser Asn Met Glu Asn Lys Ser Arg Arg Leu Ser Ser Asp Cys Tyr Leu
35 40 45

Cys Pro Lys Pro Pro Gln Thr Phe Met Leu Met Phe Tyr Xaa

50

55

60

<210> 206
<211> 44
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation

<400> 206
Met Leu Phe Leu His Thr Arg Leu His Phe Pro Arg Tyr Thr Leu Leu
1 5 10 15

Ile Cys Lys Val Leu Leu Val Val Ala Ala Ser Val His Arg Pro Trp
20 25 30

Leu Arg Ser Ile Thr Gly Cys Phe Phe Thr Lys Xaa
35 40

<210> 207
<211> 41
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals stop translation

<400> 207
Met Ser Ala Ser Leu Cys Leu Phe Thr Gln Val Leu Lys Gly Ile Val
1 5 10 15

Trp Leu Pro Ile Leu Met Phe His Val Gly Ala Thr Lys Thr Ser Gly
20 25 30

Phe Ser Val Glu Gln Leu Tyr Ser Xaa
35 40

<210> 208
<211> 57
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (57)
<223> Xaa equals stop translation

<400> 208
Met Phe Lys Arg Met Cys Phe Phe Phe Gln Val Phe Leu Pro Leu Ala
1 5 10 15

Cys Thr Glu Leu Leu Trp Lys Gly Ala Pro Cys Arg His Ile Phe Gln
20 25 30

Thr Gly Pro Asp Leu Leu Val Thr Gln Arg Cys Val His Ser Leu Leu
35 40 45

Leu Gly Tyr Leu Ile Ser Ile Phe Xaa
50 55

<210> 209

<211> 126

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (126)

<223> Xaa equals stop translation

<400> 209

Met Met Thr Gln Thr Cys Ile Ile Leu Leu Ile His Thr Met Gln Val
1 5 10 15

Cys Thr Thr His Pro Thr Val Leu Ser His Thr Leu Leu Gln Arg Pro
20 25 30

Lys Pro Thr Asp Leu Phe Pro Lys Ala Thr Pro Thr Thr Ala Pro Met
35 40 45

Pro Leu Arg Met Arg Pro Pro Gln Cys Leu Pro His Met Phe His Leu
50 55 60

Gln Ser Arg Arg Phe Asp Gln Glu Ile Gly Leu Gln Gln Lys Ser Met
65 70 75 80

Thr Gly Ile Leu Gln Thr Glu Lys Trp Thr Gln Glu Asn Phe Gly Leu
85 90 95

Ser Gln Gly Val Phe Leu Asn Met Asn Leu Ala Ser His Gln Phe Phe
100 105 110

Ser Met Lys Asp Gln Leu Pro Ser Leu Lys Leu Pro Asp Xaa
115 120 125

<210> 210

<211> 26

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (26)

<223> Xaa equals stop translation

<400> 210

Met Val Asn Ile **[REDACTED]** Gly Phe Val Ser Cys Ile Val **[REDACTED]** Val Val Ala
1 5 10 15

Val Gln Leu Cys Tyr Met Lys Gln Pro Xaa
20 25

<210> 211
<211> 48
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (48)
<223> Xaa equals stop translation

<400> 211
Met Leu Gln Phe Leu Leu Gly Phe Thr Leu Gly Asn Val Val Gly Met
1 5 10 15

Tyr Leu Ala Gln Asn Tyr Asp Ile Pro Asn Leu Ala Lys Lys Leu Glu
20 25 30

Glu Ile Lys Lys Asp Leu Asp Ala Lys Lys Lys Pro Pro Ser Ala Xaa
35 40 45

<210> 212
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation

<400> 212
Met Ala Ser Gly Ser Trp Thr Ser Ala Pro Gly Ile Gly Val Ile Leu
1 5 10 15

Val Met Thr Val Cys Leu Ser His Cys Tyr Thr His Glu Trp Gly Leu
20 25 30

Trp Gly Gly Gly Thr Gln Gly Leu Thr Asp Ser Xaa
35 40 45

<210> 213
<211> 52
<212> PRT
<213> Homo sapiens

<220>

<221> SITE

<222> (52)

<223> Xaa equals stop translation

<400> 213

Met Tyr Ile Leu Cys Ser Gly Leu Leu Gln Gly Gln Leu His Tyr Phe
1 5 10 15

Leu Gly Trp Ala Phe Leu Trp Leu Lys Leu Gly Cys Pro Trp Leu Ser
20 25 30

Gln Gly Ser Gln Pro Lys Arg His Ser Gly Glu Asn Leu Trp Pro Ile
35 40 45

Arg Glu Glu Xaa

50

<210> 214

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 214

Met Tyr Ser Leu Val Leu Thr Phe Leu Val Ser Phe Cys Ala Leu Ser
1 5 10 15

Lys Thr Phe Leu Asp His Trp Phe Gln Met Phe Ile Tyr Tyr Ile Leu
20 25 30

Phe Lys Asp Ser Glu Ile Gly Phe Cys His Pro Leu Leu Tyr Val Leu
35 40 45

Phe His Xaa

50

<210> 215

<211> 210

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (135)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (143)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE
<222> (179)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (182)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (210)
<223> Xaa equals stop translation

<400> 215

Met Arg Ser Thr Ile Leu Leu Phe Cys Leu Leu Gly Ser Thr Arg Ser
1 5 10 15

Leu Pro Gln Leu Lys Pro Ala Leu Gly Leu Pro Pro Thr Lys Leu Ala
20 25 30

Pro Asp Gln Gly Thr Leu Pro Asn Gln Gln Ser Asn Gln Val Phe
35 40 45

Pro Ser Leu Ser Leu Ile Pro Leu Thr Gln Met Leu Thr Leu Gly Pro
50 55 60

Asp Leu His Leu Leu Asn Pro Ala Ala Gly Met Thr Pro Gly Thr Gln
65 70 75 80

Thr His Pro Leu Thr Leu Gly Gly Leu Asn Val Gln Gln Gln Leu His
85 90 95

Pro His Val Leu Pro Ile Phe Val Thr Gln Leu Gly Ala Gln Gly Thr
100 105 110

Ile Leu Ser Ser Glu Glu Leu Pro Gln Ile Phe Thr Ser Leu Ile Ile
115 120 125

His Ser Leu Phe Pro Gly Xaa Ile Leu Pro Thr Ser Gln Ala Xaa Ala
130 135 140

Asn Pro Asp Val Gln Asp Gly Ser Leu Pro Ala Gly Gly Ala Gly Val
145 150 155 160

Asn Pro Ala Thr Gln Gly Thr Pro Ala Gly Arg Leu Pro Thr Pro Ser
165 170 175

Gly Thr Xaa Asp Asp Xaa Ala Val Thr Thr Pro Ala Gly Ile Gln Arg
180 185 190

Ser Thr His Ala Ile Glu Glu Ala Thr Thr Glu Ser Ala Asn Gly Ile
195 200 205

Gln Xaa
210

<210> 216
<211> 195
<212> PRT
<213> Homo sapiens

<400> 216
Met Ala Pro Ala Ala Ser Arg Leu Arg Ala Glu Ala Gly Leu Gly Ala
1 5 10 15

Leu Pro Arg Arg Ala Leu Ala Gln Tyr Leu Leu Phe Leu Arg Leu Tyr
20 . 25 30

Pro Val Leu Thr Lys Ala Ala Thr Ser Gly Ile Leu Ser Ala Leu Gly
35 40 45

Asn Phe Leu Ala Gln Met Ile Glu Lys Lys Arg Lys Lys Glu Asn Ser
50 55 60

Arg Ser Leu Asp Val Gly Gly Pro Leu Arg Tyr Ala Val Tyr Gly Phe
65 70 75 80

Phe Phe Thr Gly Pro Leu Ser His Phe Phe Tyr Phe Phe Met Glu His
85 90 95

Trp Ile Pro Pro Glu Val Pro Leu Ala Gly Leu Arg Arg Leu Leu
100 105 110

Asp Arg Leu Val Phe Ala Pro Ala Phe Leu Met Leu Phe Phe Leu Ile
115 120 125

Met Asn Phe Leu Glu Gly Lys Asp Ala Ser Ala Phe Ala Ala Lys Met
130 135 140

Arg Gly Gly Phe Trp Pro Ala Leu Arg Met Asn Trp Arg Val Trp Thr
145 150 155 160

Pro Leu Gln Phe Ile Asn Ile Asn Tyr Val Pro Leu Lys Phe Arg Val
165 170 175

Leu Phe Ala Asn Leu Ala Ala Leu Phe Trp Tyr Ala Tyr Leu Ala Ser
180 185 190

Leu Gly Lys
195

<210> 217
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 217
Met Gln Ala Arg Trp Phe His Ile Leu Gly Met Met Met Phe Ile Trp

1

5

10

15

Ser Ser Ala His Gln Tyr Lys Cys Pro Cys Tyr Ser Arg Gln Ser Gln
20 25 30

Glu Lys Xaa
35

<210> 218
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (72)
<223> Xaa equals stop translation

<400> 218
Met Phe Pro Ser Cys Leu Pro Leu Leu Phe Asn Ala Lys Val Leu Ala
1 5 10 15

Lys Asp Ile Phe Leu Leu Leu Cys Phe Ser Ile Leu Phe Cys Thr
20 25 30

Val Gly Trp Leu Ser Ala Pro Thr Leu Gly Thr Gly Pro Trp Leu Gly
35 40 45

His Phe Met Ala Gln Ser Leu Trp Gly Leu Lys Glu Gly Trp Ala Ala
50 55 60

Gln Ser Leu His Gly Ser Cys Xaa
65 70

<210> 219
<211> 53
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (53)
<223> Xaa equals stop translation

<400> 219
Met Ala Val Ser Leu Trp Pro Glu Gly Ser Gly Pro Leu Cys Ala Leu
1 5 10 15

Ser Leu Leu Thr Cys Cys Leu Val Leu Arg Pro Ala Ser Ser Gly
20 25 30

Phe Leu Trp Ser Leu Glu Glu Thr Pro Ala Leu Gln Gly Leu Cys Glu
35 40 45

Ile Ala Gln Pro Xaa
50

<210> 220
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (69)
<223> Xaa equals stop translation

<400> 220
Met Val His Asn Cys Leu Leu Leu Lys Phe Leu Leu Leu Phe Cys
1 5 10 15

Phe Pro Leu Ile Ser Tyr Gln Leu Met Asn Gly Ser Leu Gln Ser Leu
20 25 30

Gln Arg Leu Arg Met Ile Gln Asn Val Gln Cys Ile Val Leu Asn Lys
35 40 45

Gin Glu Ala Glu Phe Leu Met Gly Ile Ser Phe Gln Ile Tyr Asp Trp
50 55 60

Ser Leu Gly Phe Xaa
65

<210> 221
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (69)
<223> Xaa equals stop translation

<400> 221
Met Ser His Leu Gln Thr Leu His Leu Ile Gly Leu Ser Cys Ser Phe
1 5 10 15

Leu Tyr Phe Pro Thr Ser Gln Ala Val Glu Ala Ala Glu Pro Gly Met
20 25 30

Met Leu Ser Leu Arg Gln Met Thr Asn Pro Leu Val Ala Arg Asn Gln
35 40 45

Thr Ala Pro Arg Ala Gly Val Ser Val Phe Cys Thr Asp Cys Leu Phe
50 55 60

Gly Leu Asp Ile Xaa
65

<210> 222
<211> 44

<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation

<400> 222
Met Leu Thr Cys Ile Asp Met Asp Trp Lys Val Leu Thr Trp Leu Arg
1 5 10 15

Tyr Thr Leu Trp Ile Pro Leu Tyr Pro Leu Gly Met Phe Gly Gly Ser
20 25 30

Cys Leu Ser Asp Ser Val His Ser Asn Ile Gln Xaa
35 40

<210> 223
<211> 103
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (103)
<223> Xaa equals stop translation

<400> 223
Met Trp Ser Ser Ile Arg Leu Leu Ser Pro Val Leu Ser Leu Ile Leu
1 5 10 15

Leu Leu Ile Ala Leu Glu Leu Val Asn Ile His Ala Val Cys Gly Lys
20 25 30

Asn Ala His Glu Tyr Gln Gln Tyr Leu Lys Phe Val Lys Ser Ile Leu
35 40 45

Gln Tyr Thr Glu Asn Leu Val Ala Tyr Thr Ser Tyr Glu Lys Asn Lys
50 55 60

Trp Asn Glu Thr Ile Asn Leu Thr His Thr Ala Leu Leu Lys Met Trp
65 70 75 80

Thr Phe Ser Glu Lys Lys Gln Met Leu Ile His Leu Ala Lys Lys Ser
85 90 95

Thr Ser Lys Val Leu Leu Xaa
100

<210> 224
<211> 214
<212> PRT
<213> Homo sapiens

<220>

<221> SITE

<222> (214)

<223> Xaa equals stop translation

<400> 224

Met Lys Gly Phe Ser Trp Ala Ile Val Pro Ala Leu Thr Ser Leu Gly
1 5 10 15

Tyr Leu Ile Ile Leu Val Val Ser Ile Phe Pro Phe Trp Val Arg Leu
20 25 30

Thr Asn Glu Glu Ser His Glu Val Phe Phe Ser Gly Leu Phe Glu Asn
35 40 45

Cys Phe Asn Ala Lys Cys Trp Lys Pro Arg Pro Leu Ser Ile Tyr Ile
50 55 60

Ile Leu Gly Arg Val Phe Leu Leu Ser Ala Val Phe Leu Ala Phe Val
65 70 75 80

Thr Thr Phe Ile Met Met Pro Phe Ala Ser Glu Phe Phe Pro Arg Thr
85 90 95

Trp Lys Gln Asn Phe Val Leu Ala Cys Ile Ser Phe Phe Thr Gly Ala
100 105 110

Cys Ala Phe Leu Ala Leu Val Leu His Ala Leu Glu Ile Lys Ala Leu
115 120 125

Arg Met Lys Leu Gly Pro Leu Gln Phe Ser Val Leu Trp Pro Tyr Tyr
130 135 140

Val Leu Gly Phe Gly Ile Phe Leu Phe Ile Val Ala Gly Thr Ile Cys
145 150 155 160

Leu Ile Gln Glu Met Val Cys Pro Cys Trp His Leu Leu Ser Thr Ser
165 170 175

Gln Ser Met Glu Glu Asp His Gly Ser Leu Tyr Leu Asp Asn Leu Glu
180 185 190

Ser Leu Gly Gly Glu Pro Ser Ser Val Gln Lys Glu Thr Gln Val Thr
195 200 205

Ala Glu Thr Val Ile Xaa
210

<210> 225

<211> 109

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (34)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (48)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (109)

<223> Xaa equals stop translation

<400> 225

Met	Thr	Val	Ser	Gly	Thr	Val	Val	Leu	Val	Ala	Gly	Thr	Leu	Cys	Phe
1									10					15	

Ala	Trp	Trp	Ser	Glu	Gly	Asp	Ala	Thr	Ala	Gln	Pro	Gly	Gln	Leu	Ala
									25					30	

Pro	Xaa	Thr	Glu	Tyr	Pro	Val	Pro	Glu	Gly	Pro	Ser	Pro	Leu	Leu	Xaa
									40					45	

Ser	Val	Ser	Phe	Val	Cys	Cys	Gly	Ala	Gly	Gly	Leu	Leu	Leu	Ile	
								50			55		60		

Gly	Leu	Leu	Trp	Ser	Val	Lys	Ala	Ser	Ile	Pro	Gly	Pro	Pro	Ser	Met
								65			70		75		80

Gly	Pro	Leu	Ser	Pro	Leu	Gln	Arg	Pro	Val	Leu	Pro	His	Cys	Gly	Val
								85			90		95		

Leu	Arg	Glu	Gly	Glu	Leu	Gln	Asp	Pro	Gln	Ser	Gly	Xaa			
								100			105				

<210> 226

<211> 316

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (316)

<223> Xaa equals stop translation

<400> 226

Met	Glu	Ser	Leu	Tyr	Asp	Leu	Trp	Glu	Phe	Tyr	Leu	Pro	Tyr	Leu	Tyr
1								5			10		15		

Ser	Cys	Ile	Ser	Leu	Met	Gly	Cys	Leu	Leu	Leu	Leu	Cys	Thr	Pro	
								20				25		30	

Val	Gly	Leu	Ser	Arg	Met	Phe	Thr	Val	Met	Gly	Gln	Leu	Leu	Val	Lys
								35			40		45		

Pro	Thr	Ile	Leu	Glu	Asp	Leu	Asp	Glu	Gln	Ile	Tyr	Ile	Ile	Thr	Leu
								50			55		60		

Glu	Glu	Glu	Ala	Leu	Gln	Arg	Arg	Leu	Asn	Gly	Leu	Ser	Ser	Ser	Val
								65			70		75		80

Glu Tyr Asn Ile Met Glu Leu Glu Gln Glu Leu Glu Asn Val Lys Thr
85 90 95

Leu Lys Thr Lys Leu Asp Pro Trp Ser Ser Phe Ser Val Leu Gln Ser
100 105 110

Pro Val Trp His Phe Ala Ala Gln Thr Pro Ala Asp Ile Val Ser Pro
115 120 125

Asp Ser His Phe Met Leu Ser Thr Gln Gly Met Ser Trp Ala Gln Leu
130 135 140

Val Phe Leu Leu Pro Ala Ser Arg Pro Gly Asn Ser Gln Asp Lys Arg
145 150 155 160

Arg Lys Lys Ala Ser Ala Trp Glu Arg Asn Leu Val Tyr Pro Ala Val
165 170 175

Met Val Leu Leu Leu Ile Glu Thr Ser Ile Ser Val Leu Leu Val Ala
180 185 190

Cys Asn Ile Leu Cys Leu Leu Val Asp Glu Thr Ala Met Pro Lys Gly
195 200 205

Thr Arg Gly Pro Gly Ile Gly Asn Ala Ser Leu Ser Thr Phe Gly Phe
210 215 220

Val Gly Ala Ala Leu Glu Ile Ile Leu Ile Phe Tyr Leu Met Val Ser
225 230 235 240

Ser Val Val Gly Phe Tyr Ser Leu Arg Phe Phe Gly Asn Phe Thr Pro
245 250 255

Lys Lys Asp Asp Thr Thr Met Thr Lys Ile Ile Gly Asn Cys Val Ser
260 265 270

Ile Leu Val Leu Ser Ser Ala Leu Pro Val Met Ser Arg Thr Leu Gly
275 280 285

Leu His Lys Leu His Leu Pro Asn Thr Ser Arg Asp Ser Glu Thr Ala
290 295 300

Lys Pro Ser Val Asn Gly His Gln Lys Ala Leu Xaa
305 310 315

<210> 227

<211> 116

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (116)

<223> Xaa equals stop translation

<400> 227

Met Leu Ala Leu Ser Ser Ser Phe Leu Val Leu Ser Tyr Leu Leu Thr
 1 5 10 15

Arg Trp Cys Gly Ser Val Gly Phe Ile Leu Ala Asn Cys Phe Asn Met
 20 25 30

Gly Ile Arg Ile Thr Gln Ser Leu Cys Phe Ile His Arg Tyr Tyr Arg
 35 40 45

Arg Ala Pro Thr Gly Pro Trp Leu Ala Cys Thr Tyr Arg Gln Ser Cys
 50 55 60

Ser Gly His Leu Pro Ser Val Val Gly Leu Leu Leu Phe Arg Arg Tyr
 65 70 75 80

Ser Ser Ala Val Ser Arg Ala Gly Gln Pro Asp Trp His Thr Leu Leu
 85 90 95

Trp Gly Pro Ser Val Trp Glu Gln Leu Ser Gly Gln His Ser Ser Gln
 100 105 110

Arg Pro Ser Xaa
 115

<210> 228
 <211> 107
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (107)
 <223> Xaa equals stop translation

<400> 228

Met Cys Val Gly Trp Trp Trp Trp Leu Val Val Leu Gly Leu Gly Met
 1 5 10 15

Gly Gly Thr Leu Gly Cys Asp Gly Phe Leu Ser Gln Arg Trp Cys Phe
 20 25 30

Thr Ala Gly Lys Tyr Leu Glu Leu Gly Gly Leu Ser Arg His Gln
 35 40 45

Ala Asp Phe Ile Phe Ser Gln Thr Lys Ala Thr Phe Thr Ser Lys Gly
 50 55 60

Lys Thr Gln Asn Thr Lys Ile Glu Thr Ser Met Pro Pro His Leu Phe
 65 70 75 80

Arg Gln Gln Glu Pro Pro Gly Gln Arg Val Phe Leu Thr Leu Arg Val
 85 90 95

Thr Leu Thr Ser His Leu Val Ser Cys Gly Xaa
 100 105

<210> 229
<211> 38
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (38)
<223> Xaa equals stop translation

<400> 229
Met Ser Ser Phe Thr Leu Gly Leu Leu Phe Leu Phe Ile Phe Thr Thr
1 5 10 15

Ala Glu Asn Tyr Leu Ile Leu Phe Gln Arg Lys Tyr Cys Leu Val Ile
20 25 30
Phe Trp Gly Glu Phe Xaa
35

<210> 230
<211> 68
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (68)
<223> Xaa equals stop translation

<400> 230
Met Gln Thr Ser Gln Gln Leu Cys Cys Leu Ala Ile Ser Ile Leu Ala
1 5 10 15

Thr Leu Leu Pro Ser Gly Ala Ser Glu Glu Arg Ser Gly Leu Arg Pro
20 25 30

Gly Met Arg Leu Gln Glu Arg Glu Gln Arg Arg Ala Thr Phe Gly Ala
35 40 45

Ser Val His Ser Ser Phe Ile Ser Phe Cys Leu Leu His Gly Val Leu
50 55 60

Asn Lys Phe Xaa
65

<210> 231
<211> 51
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (51)
<223> Xaa equals stop translation

<400> 231

Met Glu Leu Ser Leu Ala Val Leu Glu Ala Val Cys Gln Cys Leu Leu
1 5 10 15

Gly Leu Trp Leu Leu Phe Trp Leu Asp Lys Glu Val Ala Val Phe Val
20 25 30

Leu Leu Leu Trp Leu Phe Thr Asp Leu Thr Asp Val Thr Gly Asp Glu
35 40 45

Cys Arg Xaa
50

<210> 232

<211> 41

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (41)

<223> Xaa equals stop translation

<400> 232

Met Lys Leu Leu Phe Cys Leu Arg Tyr Tyr Met Leu Leu Ser Val Val
1 5 10 15

Val Lys Ala Thr Ser Thr Ile Pro Ser Asn Ile Glu Ile Thr Ser Leu
20 25 30

Ser Trp Val Cys His Asn Ser Thr Xaa
35 40

<210> 233

<211> 42

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (42)

<223> Xaa equals stop translation

<400> 233

Met Arg Leu Val Ser Pro Gly Phe Trp Trp Val Leu Pro Leu Arg Leu
1 5 10 15

Gly Glu Ala Leu Pro Gly Arg Arg Gln Gln Pro Pro Gly Ala Met
20 25 30

Lys Thr Leu Arg Leu Arg Glu Val Lys Xaa
35 40

<210> 234

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 234

Met Trp Gly Pro Phe Cys Pro Phe Leu Phe Leu Ser Arg Leu Ser
1 5 10 15

Asn Ser Leu Thr Lys Asp Ser Met Asn Ile Lys Ala His Ile His Met
20 25 30

Leu Leu Glu Val Arg Ala Ala His Pro Thr Thr Arg Leu Cys Val Xaa
35 40 45

<210> 235

<211> 40

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (40)

<223> Xaa equals stop translation

<400> 235

Met Phe Ile Leu Ala Ile Trp Asn Phe Phe Ile Leu Tyr Leu Phe Ser
1 5 10 15

Thr Val Ala Gly Leu Val Cys Lys Ser Leu Cys Gln Asn Gln Thr Ile
20 25 30

Phe Lys Thr Ala Leu Cys Phe Xaa
35 40

<210> 236

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (64)

<223> Xaa equals stop translation

<400> 236

Met Leu Arg Gly Trp Ala Leu Ser Thr Phe Leu Val Cys Ile Leu Gln
1 5 10 15

Trp Val Arg Ser Leu Thr Ile Arg Leu Ala Ser Ala Leu Ser Val Arg

20

25

30

Gly Pro Ser Ser Ile Pro Ala Ser Leu Ala Ile Ile Tyr Thr Leu Phe
35 40 45

Ile Phe Ser Phe Lys Phe Leu Lys Ile Val Lys Ser Ile Tyr Ile Xaa
50 55 60

<210> 237

<211> 61

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (61)

<223> Xaa equals stop translation

<400> 237

Met Arg Lys Val Thr Ile Ser Lys Lys His Ala Leu Leu Leu Cys Phe
1 5 10 15

Gln Leu Phe Arg Cys Leu Leu Ser Met Tyr Ile Trp Ile Thr Phe Val
20 25 30

Leu Asp Gly Ser Cys Gly Ile His Cys Ser Leu Lys Pro Val Ser Phe
35 40 45

Pro Cys Thr Tyr His Ser Val His Ser Ser Thr Ser Xaa
50 55 60

<210> 238

<211> 63

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (63)

<223> Xaa equals stop translation

<400> 238

Met Cys Ala Leu Gly Val Phe Leu Leu Val Pro Trp Tyr Glu Tyr Tyr
1 5 10 15

Leu Val Leu Leu Phe Phe Pro Cys Val Ala Phe Ser Val Val Ser Gly
20 25 30

Phe Phe Leu Cys Asn Asp Ser Lys Arg Thr Leu His Ser Cys Ala Leu
35 40 45

Cys Leu Cys Ala Gly Ile Cys Phe Pro Tyr Met Phe Leu Phe Xaa
50 55 60

<210> 239

<211> 57

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (5)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (11)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (45)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (57)

<223> Xaa equals stop translation

<400> 239

Met Met Leu His Xaa Lys Leu Leu Leu Phe Xaa Glu Ala Leu Trp Tyr
1 5 10 15

Tyr Gly Gly Gly Ala Phe Leu Cys Cys Ala Gly Ser Val Pro Thr Asp
20 25 30

Cys Tyr Phe Gly Gly Leu Asp Gln Arg Arg Leu Val Xaa Asp Lys Cys
35 40 45

Thr Glu Lys Ser Thr Gly Leu Leu Xaa
50 55

<210> 240

<211> 182

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (182)

<223> Xaa equals stop translation

<400> 240

Met Thr Val Ile Leu Ile Ile Leu Ile Val Val Met Ala Arg Tyr Cys
1 5 10 15

Arg Ser Lys Asn Lys Asn Gly Tyr Glu Ala Gly Lys Lys Asp His Glu
20 25 30

Asp Phe Phe Thr Pro Gln Gln His Asp Lys Ser Lys Lys Pro Lys Lys
35 40 45

Asp Lys Lys Asn Lys Lys Ser Lys Gln Pro Leu Tyr Ser Ser Ile Val
50 55 60

Thr Val Glu Ala Ser Lys Pro Asn Gly Gln Arg Tyr Asp Ser Val Asn
65 70 75 80

Glu Lys Leu Ser Asp Ser Pro Ser Met Gly Arg Tyr Arg Ser Val Asn
85 90 95

Gly Gly Pro Gly Ser Pro Asp Leu Ala Arg His Tyr Lys Ser Ser Ser
100 105 110

Pro Leu Pro Thr Val Gln Leu His Pro Gln Ser Pro Thr Ala Gly Lys
115 120 125

Lys His Gln Ala Val Gln Asp Leu Pro Pro Ala Asn Thr Phe Val Gly
130 135 140

Ala Gly Asp Asn Ile Ser Ile Gly Ser Asp His Cys Ser Glu Tyr Ser
145 150 155 160

Cys Gln Thr Asn Asn Lys Tyr Ser Lys Gln Met Arg Leu His Pro Tyr
165 170 175

Ile Thr Val Phe Gly Xaa
180

<210> 241

<211> 71

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (71)

<223> Xaa equals stop translation

<400> 241

Met His Met Tyr Val Trp Val Arg Ala His Leu Val Phe Tyr Leu Phe
1 5 10 15

Val Cys Leu Ser Glu Ser Ser Ala Gly Gln Arg Leu Pro Leu Asp Cys
20 25 30

Cys Cys Ser Gly Asp Glu Lys Asp Glu Glu Ser Ala Gly Lys Arg Gly
35 40 45

Gly Val Gln Glu His Gly Gly His Leu Gly Pro Ser Phe Trp His Thr
50 55 60

Lys Pro Glu Phe Ser Cys Xaa
65 70

<210> 242
<211> 62
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (62)
<223> Xaa equals stop translation

<400> 242
Met Trp Arg Val Met Leu Ala Trp Leu Ala Met Val Asn Ser Pro Met
1 5 10 15

Ala Met Glu Ser Gln Val Gly His Ile Ile Ala Val Lys Asp Thr Leu
20 25 30

Thr Gln Met Thr Leu Pro Gly Ala Arg Ile Glu Pro Val Arg Lys Glu
35 40 45

Ser Lys Ala Gly Ser Ala Gly Lys Arg Glu Gly Phe Cys Xaa
50 55 60

<210> 243
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 243
Met Ile Ala Asp Trp Met Phe Phe Val Tyr Ala Leu Cys Ile Asp Val
1 5 10 15

Thr Ala Asn Glu Phe Cys Leu Thr Leu Thr Phe Leu Thr Ser Lys Val
20 25 30

Ser Lys Xaa
35

<210> 244
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation

<400> 244
Met Glu Pro Val Ala Leu Leu Gln Pro Thr Trp Trp Leu Leu Asn Val
1 5 10 15

Thr Leu Pro Leu Val Ala Trp Ser Gly Pro Leu Ile Cys Arg Pro Leu
20 25 30

Leu His Gly Glu Gly Arg Gln Gly Ala Ala Cys Leu Gln Gly Xaa
35 40 45

<210> 245

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 245

Met His Phe Lys Arg Thr Gln Asn His Leu Asn Ile Val Thr Trp Leu
1 5 10 15

Leu Gln Val Met Ile Ile Val Met Leu Ile Ile Met Arg Ile Ser Cys
20 25 30

Thr His Gln Pro Val Glu Ser Lys Lys Phe Pro Phe Arg Asn Phe Leu
35 40 45

Ser Cys Xaa

50

<210> 246

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 246

Met Thr Tyr His Val Val Cys Ala Phe Leu Ile Val Val Leu Lys Lys
1 5 10 15

Gln Phe Ile Leu Ala Leu Gln Thr Ile Ser Thr Ser Leu Arg Ser Lys
20 25 30

Gln Ile Leu Met Val Leu Ser Ser Thr Ile Ile Ala Asp Ser Thr Phe
35 40 45

Tyr Tyr Xaa

50

<210> 247

<211> 33

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (33)

<223> Xaa equals stop translation

<400> 247

Met Pro Val Pro Leu Trp Leu Val Leu Trp Phe Cys Phe Leu Leu Tyr
1 5 10 15

Val Ala Ser Arg Arg Thr Phe Gly Leu Ala Asn Tyr Met Pro Leu Pro
20 25 30

Xaa

<210> 248

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals stop translation

<400> 248

Met Leu Ile Cys Arg Leu Val Leu Leu Ala Asp Pro Gly Pro Val Asn
1 5 10 15

Phe Met Val Arg Leu Phe Val Val Ile Val Met Phe Ala Trp Ser Ile
20 25 30

Val Gly Lys Tyr Val Leu Ile Ser Thr Ile Thr Glu Gln Thr Lys Thr
35 40 45

Xaa

<210> 249

<211> 116

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (116)

<223> Xaa equals stop translation

<400> 249

Met Ile Asn Val Tyr Phe Ser Gly Pro Gly Val Leu Thr Pro Leu Asp
1 5 10 15

Asp Gln Gly Ser Pro Cys Pro Pro Ala Pro Phe Ala Ala Leu His Pro

20

25

30

Cys Pro His Pro Ala Gly Ser Gly Val Leu Cys Cys Cys Pro Leu Arg
35 40 45

Leu Cys Arg Pro Cys Arg Ile Leu Phe Thr Gly Pro Leu Leu Leu Thr
50 55 60

Leu His His Leu Leu Cys Glu Thr Ser Pro Ser Gly Ile Gly Val Gly
55 70 75 80

Asn Ile Val Pro Gly Ala Arg Pro Leu Gly Val Asn Pro Val Phe Pro
85 90 95

Ile Ser Ser Cys Asp Leu Gly Gln Val Ala Glu Pro Leu Leu Val Thr
100 105 110

Ile Ser Ser Xaa
115

<210> 250

<211> 75

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (75)

<223> Xaa equals stop translation

<400> 250

Met Thr Asn Val Tyr Ser Leu Asp Gly Ile Leu Val Phe Gly Leu Leu
1 5 10 15

Phe Val Cys Thr Cys Ala Tyr Phe Lys Lys Val Pro Arg Leu Lys Thr
20 25 30

Trp Leu Leu Ser Glu Lys Lys Gly Val Trp Gly Val Phe Tyr Lys Ala
35 40 45

Ala Val Ile Gly Thr Arg Leu His Ala Ala Val Ala Ile Ala Cys Val
50 55 60

Val Met Ala Phe Tyr Val Leu Phe Ile Lys Xaa
65 70 75

<210> 251

<211> 63

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (57)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (63)
<223> Xaa equals stop translation

<400> 251
Met Pro Thr Leu Arg Val Pro Val Leu Ser Val Trp Leu Leu Arg Trp
1 5 10 15

Trp Arg Val Leu Gly Ala Gly Arg Val Leu Pro Asp Ser Leu Ser Leu
20 25 30

Ser Pro Pro Pro Pro Thr Gly Cys Gln Thr Lys Pro Glu Arg Gly Trp
35 40 45

Gly Ser Gln Pro Pro Ser Val Leu Xaa Pro Gln Ala Pro Val Xaa
50 55 60

<210> 252
<211> 73
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (73)
<223> Xaa equals stop translation

<400> 252
Met Val Tyr Tyr Leu Asn Arg Ala Leu Arg Ala Thr Phe Ser Ile Leu
1 5 10 15

Phe Ser Val Val Cys Leu Leu Phe Leu Gly Ser Ile Val Asn Cys Phe
20 25 30

Leu Asn Asp Val Phe Lys Pro Leu Thr Leu Asn Phe Ser Thr Ala Leu
35 40 45

Ser Ala Trp Arg Lys Glu Ser Ser Ala Trp Asn Ser Leu Gly Leu Leu
50 55 60

Pro Pro Thr Asp Glu Tyr Pro Thr Xaa
65 70

<210> 253
<211> 49
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (49)
<223> Xaa equals stop translation

<400> 253
Met Val Val Asn Asp Arg Leu Val Ser Thr Cys Ile Leu Cys Thr Leu

1

5

10

15

His Ile Pro Leu Phe Phe Leu Ile Phe Leu Val Tyr Glu Val His Leu
20 25 30

Val Phe Gln Ile Val Ala Asn Leu Gln Lys Ile Phe Gln Tyr Ile Tyr
35 40 45

Xaa

<210> 254

<211> 41

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (41)

<223> Xaa equals stop translation

<400> 254

Met Ile Ile Leu His Ile Val Val Cys Leu Phe Thr Ile Ser Ile Ile
1 5 10 15

Glu Glu Gln Lys Glu Glu Ile Leu Cys Ser Thr Lys Ser Gln Ala Glu
20 25 30

Lys Thr Val Thr His Ile Glu Gln Xaa
35 40

<210> 255

<211> 54

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (54)

<223> Xaa equals stop translation

<400> 255

Met Thr Leu Ser Val Leu Phe Ala Phe Pro Ile Trp Leu Lys Tyr Leu
1 5 10 15

Asn Leu Asn Ile Phe Phe Leu Ala Leu Lys Ile Phe Trp Val Ile Leu
20 25 30

Ser Phe Cys Thr Ser Cys Thr Ser Trp Tyr Ser Gly Ala Arg Val Ile
35 40 45

Phe Phe Gln Ile Ile Xaa
50

<210> 256

<211> 41
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals stop translation

<400> 256
Met Cys Arg Arg Ile Gln Arg Leu Arg Ala Met Leu His Met Leu Leu
1 5 10 15

Val Ser Met Leu Pro Thr Val Gly Lys Pro Asn Met Tyr Gln Pro Pro
20 25 30
Gln Asn Tyr Asp Ile Leu Leu Gln Xaa
35 40

<210> 257
<211> 42
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 257
Met Ala Leu Ala Phe Leu His Leu Asn Ile Ser Xaa Ser Gln Ala Leu
1 5 10 15

Thr Leu Cys Lys Glu Leu Glu Lys Pro Lys Leu Glu Lys Asn Lys Gly
20 25 30
Gly Pro Ala Leu Glu Lys Leu Val Val Xaa
35 40

<210> 258
<211> 53
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (53)
<223> Xaa equals stop translation

<400> 258
Met Ser Gly Thr Thr Trp Thr Ala Ile His Leu Thr Ser Asn Leu Phe

1

5

10

15

Gly Ile Leu Ala Leu Pro Gly Asn Gln Ser Ser Gly Ser Asn Ile Glu
20 25 30

Gln Leu Cys Thr Ser Ser Arg Glu Ala Thr Asn Arg Leu Pro Cys Val
35 40 45

Asp Val Gly Ser Xaa
50

<210> 259

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 259

Met Phe Tyr Pro Pro Cys Pro Phe Phe Pro Gln Leu Cys Phe Cys Ile
1 5 10 15

Phe Phe Leu Gly Lys Cys Lys Leu Ser Leu Ser Phe Met Thr Cys Glu
20 25 30

Ile Ser Val Ser Leu Glu Phe Val Arg Arg Arg Gly Asn His Ala Xaa
35 40 45

<210> 260

<211> 53

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (53)

<223> Xaa equals stop translation

<400> 260

Met Asn Ser Trp Ile Leu Asn Met Arg Val Arg Phe Thr Phe Leu Ser
1 5 10 15

Gln Leu Leu Thr Leu Ile Pro Arg Thr Ser His Ser Ala Thr Ser Val
20 25 30

Gly Asn Ser Gln Ile Glu Leu Pro Arg Glu Lys His His Met Thr Tyr
35 40 45

Trp Glu Asn Gly Xaa
50

<210> 261
<211> 55
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (55)
<223> Xaa equals stop translation

<400> 261
Met Phe Ile Val Ile Cys Lys Ile Leu Leu Phe Leu Ile Leu Val Ala
1 5 10 15

Arg Pro Phe Arg Thr His Ser Cys Ile Lys Tyr Phe Ala Leu Phe Lys
20 25 30

Glu Thr His Met Asp Glu Val Arg Met Cys Asn Met Met Ala Ser Gln
35 40 45

Cys Ser Ser Leu Tyr Leu Xaa
50 55

<210> 262
<211> 38
<212> PRT
<213> Homo sapiens

<400> 262
Met Lys Asn Met Asn Ser Arg Tyr Tyr Leu Arg Ala Ile Phe Cys Leu
1 5 10 15

Tyr Thr Leu Ala Cys Ile Leu Phe Leu Gln Ile Ile Leu Lys Ala Arg
20 25 30

Cys Gly Gly Ser Arg Leu
35

<210> 263
<211> 24
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (24)
<223> Xaa equals stop translation

<400> 263
Met Pro Pro Leu Phe Leu Gly Ser Phe Leu Val Leu Trp Leu Gly Gly
1 5 10 15

Val Val Leu Cys Thr Gly Gly Xaa
20

<210> 264
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (11)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation

<400> 264
Met Val Cys Ala Leu Gly Val Tyr Val Cys Xaa Ser Ala Pro Thr Ala
1 5 10 15

Ala Val Pro Lys Pro Ala Lys Gly Thr Ile Cys Leu Lys Met Leu Ser
20 25 30

Gly Ala Asn Cys Ala Cys Gln Gly Gln Val Thr Arg Gln His Xaa
35 40 45

<210> 265
<211> 115
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (115)
<223> Xaa equals stop translation

<400> 265
Met Ala Gly Pro Arg Ala Ser Thr Gly Pro Arg Pro Xaa Cys Leu Val
1 5 10 15

Leu Phe Leu Phe Asn Phe Ile Phe Cys Phe Met Ser Val Cys Pro Pro
20 25 30

Thr Pro Thr Pro Phe Ser Val Lys Trp Gly Ala Leu Gly Glu Ser Leu
35 40 45

Leu Pro Pro Ser Leu Ser Gln Asp Leu Pro Pro Arg His Gln Pro Ser
50 55 60

Leu Trp Thr Arg Gln Arg Ala Asp Arg Val Gly Arg Gly Leu Arg Val
65 70 75 80

Ala Arg Ala Ser Pro Pro Ala Asn Gly Pro Leu Leu Arg Pro Pro Val
85 90 95

Ser Pro Cys Pro Phe Leu Lys Gln Asn Ala Leu Val Cys Lys Pro Leu
100 105 110

Asp Ala Xaa
115

<210> 266

<211> 248

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (166)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (248)

<223> Xaa equals stop translation

<400> 266

Met His Leu Ala Arg Leu Val Gly Ser Cys Ser Leu Leu Leu Leu
1 5 10 15

Gly Ala Leu Ser Gly Trp Ala Ala Ser Asp Asp Pro Ile Glu Lys Val
20 25 30

Ile Glu Gly Ile Asn Arg Gly Leu Ser Asn Ala Glu Arg Glu Val Gly
35 40 45

Lys Ala Leu Asp Gly Ile Asn Ser Gly Ile Thr His Ala Gly Arg Glu
50 55 60

Val Glu Lys Val Phe Asn Gly Leu Ser Asn Met Gly Ser His Thr Gly
65 70 75 80

Lys Glu Leu Asp Lys Gly Val Gln Gly Leu Asn His Gly Met Asp Lys
85 90 95

Val Ala His Glu Ile Asn His Gly Ile Gly Gln Ala Gly Lys Glu Ala
100 105 110

Glu Lys Leu Gly His Gly Val Asn Asn Ala Ala Gly Gln Ala Gly Lys
115 120 125

Glu Ala Asp Lys Ala Val Gln Gly Phe His Thr Gly Val His Gln Ala
130 135 140

Gly Lys Glu Ala Glu Lys Leu Gly Gln Gly Val Asn His Ala Ala Asp
145 150 155 160

Gln Ala Gly Lys Glu Xaa Glu Lys Leu Gly Pro Ser Ala His His Ala

165

170

175

Ala Gly Gln Ala Gly Lys Glu Leu Gln Asn Ala His Asn Gly Val Asn
180 185 190

Gln Ala Ser Lys Glu Ala Asn Gln Leu Leu Asn Gly Asn His Gln Ser
195 200 205

Gly Ser Ser Ser His Gln Gly Gly Ala Thr Thr Thr Pro Leu Ala Ser
210 215 220

Gly Ala Ser Val Asn Thr Pro Phe Ile Asn Leu Pro Ala Leu Trp Arg
225 230 235 240

Ser Val Ala Asn Ile Met Pro Xaa
245

<210> 267

<211> 178

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (155)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (178)

<223> Xaa equals stop translation

<400> 267

Met Leu Phe Leu Phe Leu Tyr Cys Leu Leu Val Val Leu Pro Phe Lys
1 5 10 15

Leu Thr Pro Lys His Ser Ala Glu Val Leu Leu Ser Ile His Lys Ser
20 25 30

Lys Lys Tyr Leu Cys Lys Val Lys Ala Ala Cys Lys Ile Gln Ala Trp
35 40 45

Tyr Arg Cys Trp Arg Ala His Lys Glu Tyr Leu Ala Ile Leu Lys Ala
50 55 60

Val Lys Ile Ile Gln Gly Cys Phe Tyr Thr Lys Leu Glu Arg Thr Arg
65 70 75 80

Phe Leu Asn Val Arg Ala Ser Ala Ile Ile Gln Arg Lys Trp Arg
85 90 95

Ala Ile Leu Pro Ala Lys Ile Ala His Glu His Phe Leu Met Ile Lys
100 105 110

Arg His Arg Ala Ala Cys Leu Ile Gln Ala His Tyr Arg Gly Tyr Lys
115 120 125

Gly Arg Gln Val Phe Leu Arg Gln Lys Ser Ala Ala Leu Ile Ile Gln
130 135 140

Lys Tyr Ile Arg Ala Arg Glu Ala Gly Lys Xaa Glu Arg Ile Lys Tyr
145 150 155 160

Ile Glu Phe Lys Asn Leu Gln Leu Ser Tyr Lys His Trp Cys Val Val
165 170 175

Gly Xaa

<210> 268

<211> 79

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (79)

<223> Xaa equals stop translation

<400> 268

Met Arg Pro Leu Leu Gly Leu Leu Val Phe Ala Gly Cys Thr Phe
1 5 10 15

Ala Leu Tyr Leu Leu Ser Thr Arg Leu Pro Arg Gly Arg Arg Leu Gly
20 25 30

Ser Thr Glu Glu Ala Gly Gly Arg Ser Leu Trp Phe Pro Ser Asp Leu
35 40 45

Ala Glu Leu Arg Glu Leu Ser Glu Val Leu Arg Glu Tyr Arg Lys Glu
50 55 60

His Gln Ala Tyr Val Phe Leu Leu Phe Cys Gly Ala Tyr Leu Xaa
65 70 75

<210> 269

<211> 81

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (81)

<223> Xaa equals stop translation

<400> 269

Met Lys Leu Ser Gly Met Phe Leu Leu Ser Leu Ala Leu Phe Cys
1 5 10 15

Phe Leu Thr Gly Val Phe Ser Gln Gly Gln Val Asp Cys Gly Glu
20 25 30

Phe Gln Asp Thr Lys Val Tyr Cys Thr Arg Glu Ser Asn Pro His Cys

35

40

45

Gly Ser Asp Gly Gln Thr Tyr Gly Asn Lys Cys Ala Phe Cys Lys Ala
50 55 60

Ile Val Lys Ser Gly Gly Lys Ile Ser Leu Lys His Pro Gly Lys Cys
65 70 75 80

Xaa

<210> 270

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (69)

<223> Xaa equals stop translation

<400> 270

Met Asp Ala Ala Met Pro Val Cys Pro Cys Leu Ile Cys Val Cys Phe
1 5 10 15

Val Leu Arg Leu Gln Ser Gly Val Ala Gly Thr Glu Thr Glu Arg Pro
20 25 30

Pro His Gly Ala Ala Ser Leu His Gln Asp Arg Gly Ala Thr Leu Arg
35 40 45

Leu Cys Phe Phe Pro Ser Gly Val Gly Phe Leu Leu Phe Leu Ser Ile
50 55 60

Leu Pro Trp Ser Xaa
65

<210> 271

<211> 131

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (131)

<223> Xaa equals stop translation

<400> 271

Met Asn Phe Arg Gln Arg Met Gly Trp Ile Gly Val Gly Leu Tyr Leu
1 5 10 15

Leu Ala Ser Ala Ala Ala Phe Tyr Tyr Val Phe Glu Ile Ser Glu Thr
20 25 30

Tyr Asn Arg Leu Ala Leu Glu His Ile Gln Gln His Pro Glu Glu Pro
35 40 45

Leu Glu Gly Thr Thr Trp Thr His Ser Leu Lys Ala Gln Leu Leu Ser
50 55 60

Leu Pro Phe Trp Val Trp Thr Val Ile Phe Leu Val Pro Tyr Leu Gln
65 70 75 80

Met Phe Leu Phe Leu Tyr Ser Cys Thr Arg Ala Asp Pro Lys Thr Val
85 90 95

Gly Tyr Cys Ile Ile Pro Ile Cys Leu Ala Val Ile Cys Asn Arg His
100 105 110

Gln Ala Phe Val Lys Ala Ser Asn Gln Ile Ser Arg Leu Gln Leu Ile
115 120 125

Asp Thr Xaa
130

<210> 272

<211> 85

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (65)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (85)

<223> Xaa equals stop translation

<400> 272

Met Trp Val Phe Phe Leu Pro Phe Phe Ser Ile Leu Phe Lys Ile Cys
1 5 10 15

Trp Cys Ile Ser Leu Ser Gln Thr Lys Glu Lys Gln Ser Ser Asn Leu
20 25 30

Met Phe Tyr Phe Phe Cys Ile Cys Thr Tyr Glu Arg Arg Arg Lys Lys
35 40 45

Glu Met Arg Arg Gly Glu Lys Lys Arg Ser Phe Cys Leu Ile Gly Leu
50 55 60

Xaa Gln His Met Ile Ala Val Gln Ala Trp Phe His Glu Gln His Gln
65 70 75 80

Ile Gln Ile Ser Xaa
85

<210> 273

<211> 79

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (61)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (79)

<223> Xaa equals stop translation

<400> 273

Met Gln Trp Pro Phe Leu Cys Val Leu Pro Leu Leu Pro Gln Val Trp
1 5 10 15

Arg Ala Gly Ser Leu Leu Arg Ala Leu Glu Leu Tyr Ser Val Leu Leu
20 25 30

Ser His Phe Leu Trp Glu Met Trp Thr Met Ser Leu Lys Glu Pro Glu
35 40 45

Leu Leu Leu Ser Thr Lys Ser Leu Thr Val Trp Arg Xaa Arg Glu Pro
50 55 60

Leu Ser Glu Ile Gly Gly Cys Arg Leu Asn Asn Glu Gly Thr Xaa
65 70 75

<210> 274

<211> 54

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (54)

<223> Xaa equals stop translation

<400> 274

Met Phe Cys Phe Asn Trp Leu Leu Cys Phe Leu Phe Pro Arg Phe Pro
1 5 10 15

Ile Leu Val Cys Arg Lys His Gln Phe Cys Val Tyr Leu Leu Leu Val
20 25 30

Leu Lys Leu Arg Thr Leu Tyr Ala Glu Leu Ile Asp Leu His Leu Cys
35 40 45

Ala Ser Ile Leu Gly Xaa

50

<210> 275

<211> 155

<212> PRT

<213> Homo sapiens

<220>
<221> SITE
<222> (150)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (155)
<223> Xaa equals stop translation

<400> 275
Met Ala Arg His Gly Leu Pro Leu Leu Pro Leu Leu Ser Leu Leu Val
1 5 10 15

Gly Ala Trp Leu Lys Leu Gly Asn Gly Gln Ala Thr Ser Met Val Gln
20 25 30

Leu Gln Gly Gly Arg Phe Leu Met Gly Thr Asn Ser Pro Asp Ser Arg
35 40 45

Asp Gly Glu Gly Pro Val Arg Glu Ala Thr Val Lys Pro Phe Ala Ile
50 55 60

Asp Ile Phe Pro Val Thr Asn Lys Asp Phe Arg Asp Phe Val Arg Glu
65 70 75 80

Lys Lys Tyr Arg Thr Glu Ala Glu Met Phe Gly Trp Ser Phe Val Phe
85 90 95

Glu Asp Phe Val Ser Asp Glu Leu Arg Asn Lys Ala Thr Gln Pro Met
100 105 110

Lys Ser Val Leu Trp Trp Leu Pro Val Glu Lys Ala Phe Trp Arg Gln
115 120 125

Pro Ala Gly Pro Gly Ser Gly Ile Arg Glu Arg Leu Glu His Pro Val
130 135 140

Leu His Val Ser Trp Xaa Asp Ala Arg Ala Xaa
145 150 155

<210> 276
<211> 129
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (68)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (98)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE
<222> (103)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (104)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (112)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (114)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (124)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (129)
<223> Xaa equals stop translation

<400> 276
Met Ala Tyr Arg His Phe Trp Met Leu Val Leu Phe Val Ile Phe Asn
1 5 10 15

Ser Leu Gln Gly Leu Tyr Val Phe Met Val Tyr Phe Ile Leu His Asn
20 25 30

Gln Met Cys Cys Pro Met Lys Ala Ser Tyr Thr Val Glu Met Asn Gly
35 40 45

His Pro Gly Pro Ser Thr Ala Phe Phe Thr Pro Gly Ser Gly Met Pro
50 55 60

Pro Ala Gly Xaa Glu Ile Ser Lys Ser Thr Gln Asn Leu Asn Arg Trp
65 70 75 80

Tyr Gly Gly Arg Cys His Leu Thr Gly Arg Glu His Pro Ser Lys Gln
85 90 95

Gly Xaa Gln Gly Gln Pro Xaa Xaa Lys Ala Lys Ser Thr Lys Trp Xaa
100 105 110

His Xaa Pro Val Leu Trp Arg Ile Trp Pro Gly Xaa Thr Asp Ser Arg
115 120 125

Xaa

<210> 277
<211> 84
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (84)
<223> Xaa equals stop translation

<400> 277
Met Ala Ser Pro Gly Trp His Leu Ser Cys Arg Pro Thr Gly Leu Val
1 5 10 15

Ser Ile Phe Leu Leu Cys Ala Pro Ala Tyr Leu His Ser Phe Val Met
20 25 30

Thr Ser Ile Thr Leu Ile Ser Thr Lys Ile Cys Ser Pro Thr Lys Leu
35 40 45

Arg His Arg Thr His Phe Leu Tyr Gly Ser Ile Met Glu Leu Tyr Pro
50 55 60

Thr Leu Thr Phe Pro Met Thr Thr Asp Val Glu Asn Leu Asn Leu Asp
65 70 75 80

Ser Ser Arg Xaa

<210> 278
<211> 86
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (86)
<223> Xaa equals stop translation

<400> 278
Met Gly Cys Arg Gly Asn Lys Leu Phe Val Leu Ser Tyr Cys Thr Cys
1 5 10 15

Leu Thr Trp Leu Leu Gly Thr Lys Ser Gln Lys Asn Pro Phe Gln Val
20 25 30

Cys Met Ser Gly Gly Trp Ala Val Ser Arg Leu Glu Thr Gly Phe Gln
35 40 45

Ala Leu His Asp Gly Arg Ala Ser Ser Pro Leu Ser Ala Ala Cys Val
50 55 60

Leu Asp Arg Thr Val Ala Arg Arg Trp Lys Pro Pro Ser Val Pro Leu
65 70 75 80

Ala His His Thr Lys Xaa
85

<210> 279
<211> 96
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (96)
<223> Xaa equals stop translation

<400> 279
Met Pro Trp Leu Thr Ile Leu Arg Phe Leu Gln Ala Ser Gly His Val
1 5 10 15

Arg Ala Gln Asp Leu Ala Leu Leu Gly Asp Thr Ser Val Cys Ile Arg
20 25 30

Cys Gly Cys Gly Cys Ser Leu Ser Ile Ala Asn Tyr Glu Trp Val
35 40 45

Pro Leu Arg Arg Lys Asp Cys Lys Arg Tyr Glu Thr Ser Glu Lys Thr
50 55 60

Ser Cys Leu Leu Leu Pro Ser Ala Cys Ser Arg Gln Asn Ala Val Gly
65 70 75 80

Phe Ser Arg Leu Pro Val Pro Lys Leu Ser Cys Leu Leu His Gly Xaa
85 90 95

<210> 280
<211> 98
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (70)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (98)
<223> Xaa equals stop translation

<400> 280
Met Ile Leu Leu Phe Leu Leu Ser Leu Ser Leu Ser Leu Ser Leu
1 5 10 15

Ser Leu Ser Phe Ser Pro Leu Asn Cys Leu Phe Ser Phe Trp Gly Ser
20 25 30

Pro Pro Thr Arg Cys Ser Trp Cys Arg Leu Gly Ser Gln Gly Glu Ala

35

40

45

Trp Trp Pro Gly Leu Gly Arg Gly Thr Leu Ser Leu Ala Lys Ala Glu
50 55 60

Ser Glu Ile Val Val Xaa Leu Cys Lys Ser Tyr Phe Gln Tyr Phe Leu
65 70 75 80

Ala Ala Ser Glu Val Ser Leu Thr Pro Cys Arg Ala Leu Leu Leu
85 90 95

Ser Xaa

<210> 281

<211> 55

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (55)

<223> Xaa equals stop translation

<400> 281

Met Ser Val Trp Pro Arg Ser Thr Leu Leu Phe Cys Leu Leu Ser Leu
1 5 10 15

Ser Thr Gly Leu Phe Leu Asp Lys Leu Gly Ile Ile Ile Pro Ile Leu
20 25 30

Leu Cys Gly Trp Lys Leu Asn Val Ile Met Met Cys Val Arg Cys Leu
35 40 45

His Ser Ala Trp Arg Tyr Xaa
50 55

<210> 282

<211> 72

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (72)

<223> Xaa equals stop translation

<400> 282

Met Arg Ile His Phe Lys Ile Leu Val Leu Val Ile Tyr Phe Ile Leu
1 5 10 15

Leu Gly Ser Phe Ser Asp Arg Cys Ser Leu Leu Asp Cys Lys Ser Arg
20 25 30

Ile Gln Arg Ile Phe Ile Cys Asn Ile Leu Asn Leu Ser Leu Val Ser
35 40 45

Cys His Leu Cys Arg Tyr Ser Phe Asp Cys Leu Thr Arg Gly Lys Cys
50 55 60

Phe Pro Leu Ser Phe Pro Ala Xaa
65 70

<210> 283
<211> 44
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation

<400> 283
Met Tyr Ala Ala Ala Leu Ser Thr Ala Pro Ser Leu Phe Phe Leu His
1 5 10 15

Leu Cys Leu Leu Lys Thr Leu Ile Leu Phe Ser Leu Ser Ser Ile Pro
20 25 30

Leu Pro Pro Leu Leu Tyr Ser Tyr Asp Leu His Xaa
35 40

<210> 284
<211> 56
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation

<400> 284
Met Leu Pro Ser Asn Trp Ser Gly Thr Trp Ala Leu Ile Gln Leu Ser
1 5 10 15

Ile Pro Phe Thr Leu Ala Phe His Gln Pro Asn Lys Asn Gln Leu Thr
20 25 30

Gln Lys Lys Arg Lys Ala Pro Gln Gly Ser Phe Asp Pro Asp Ile Tyr
35 40 45

Ile Asp Ala Ile Gly Val Pro Xaa
50 55

<210> 285
<211> 49
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (49)
<223> Xaa equals stop translation

<400> 285

Met Ser Thr Leu Arg Arg Met Ala Leu Leu Tyr Ile Glu Thr Pro Leu
1 5 10 15

Leu Arg Ala Leu Met Val Gln Gly Pro Arg Leu Val Ser Val Arg Ala
20 25 30

Ala Met His Gly Lys Cys Gly Gly Arg Ala Leu Trp Ala Leu Trp Gln
35 40 45

Xaa

<210> 286
<211> 42
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 286

Met Val Cys Val Arg Cys Val Trp Tyr Val Trp His Val Phe Gly Val
1 5 10 15

Tyr Gly Asn Ile Leu Trp Ile Arg Thr Cys Gly Leu Phe Lys Asp Leu
20 25 30

Ser Phe Cys Ala Leu Lys Ser Glu Met Xaa
35 40

<210> 287
<211> 49
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (49)
<223> Xaa equals stop translation

<400> 287

Met Arg His Val Ala Ile Val Thr Met Ile Val Val Leu Ser Pro Pro
1 5 10 15

Val Leu Ala Ser Ser Leu Lys Pro Pro Leu Phe Ile Asp Thr Tyr Phe
20 25 30

Met Phe Gly Lys Arg Cys Ser Arg Trp Asp Thr Pro Ala Cys Ser Lys

35

40

45

Xaa

<210> 288

<211> 110

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (110)

<223> Xaa equals stop translation

<400> 288

Met	Trp	Ala	Glu	Leu	Lys	Leu	Leu	Ser	Trp	Gly	Arg	Ala	Ala	Ile	Ala
1				5					10					15	

Val	Trp	Val	Cys	Leu	Arg	Arg	Val	Val	Arg	Gly	Gly	His	Ser	Pro	Pro
				20				25					30		

Ala	Gly	Gln	Gly	Gly	Gln	Gly	Val	Lys	Val	Gln	Trp	Glu	Gly	Val	Gln
		35					40					45			

Gly	Ser	Gly	Ser	Gly	Gln	Pro	Glu	Asp	Met	Arg	Trp	Glu	Lys	Leu	His
	50					55				60					

Val	Arg	Ile	Leu	Met	Gln	Gly	Met	His	Gly	Ala	Pro	Gln	Asp	Asp	Ile
	65				70				75				80		

Arg	Ser	Val	His	Gly	Ser	Thr	Ala	Phe	Pro	Asp	Cys	Leu	His	Leu	Pro
	85					90						95			

Cys	Arg	Pro	Thr	Cys	Pro	Gly	Val	Ser	Phe	Gly	Ser	Gly	Xaa		
				100				105				110			

<210> 289

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (64)

<223> Xaa equals stop translation

<400> 289

Met	Leu	Leu	Val	Ser	Cys	Phe	Met	Ser	Ile	Tyr	Phe	Leu	Ser	Pro	Leu
1				5					10				15		

Leu	Leu	Pro	Leu	His	Gly	Ser	Pro	His	Pro	His	Ser	Tyr	Leu	Cys	Phe
				20				25				30			

Ala	Val	Cys	Arg	Thr	Ser	Trp	Ser	Leu	Ser	Glu	Lys	Thr	Cys	Asn	Phe
			35				40				45				

Pro Asn Glu Met Leu Gln Leu Pro Ile Phe Leu Lys Ser Ile Tyr Xaa
50 55 60

<210> 290
<211> 42
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 290
Met Gly Leu Leu Leu Leu Leu Gly Cys Trp Thr His Ile Phe
1 5 10 15

Phe Thr Asn Gly Met Ile Tyr Trp Tyr Leu Glu Gly His Pro Ile Leu
20 25 30

Asn Glu Ile Leu Phe Ile Leu His Phe Xaa
35 40

<210> 291
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation

<400> 291
Met Ile Asn Cys Val Cys Val His Ala Cys Val Arg Ala Cys Gly Leu
1 5 10 15

Leu His Ser Leu Val Leu Leu Ser Leu Ser Leu Ser Ala Leu
20 25 30

Phe Ile Pro Trp Asp Thr Glu Ile Phe Lys Xaa
35 40

<210> 292
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)

<223> Xaa equals stop translation

<400> 292

Met Leu Phe Phe Cys Leu Leu Met Lys Met Leu Gly Pro Ser Arg Leu
1 5 10 15

Pro Phe Leu Ala Leu Thr Leu Cys Arg Phe Ile Leu Tyr Phe Gln Phe
20 25 30

Cys Tyr Leu Ile Ser Asp Ser Ser Pro Asp His Ser Xaa
35 40 45

<210> 293

<211> 57

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (57)

<223> Xaa equals stop translation

<400> 293

Met Cys Phe Thr Gln Phe Ser Arg Ile Phe Phe Leu Thr Ser Ser Leu
1 5 10 15

Thr Leu Ala Ala Cys Ala Asn His Ile Leu Ala Ala Tyr Ser Ser Ser
20 25 30

Leu Ala Asp Arg Cys Val Gly Glu Lys Ser Leu Ile Val Ile Val Pro
35 40 45

Glu Arg Ser Phe Gln Thr His Phe Xaa

50 55

<210> 294

<211> 75

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (75)

<223> Xaa equals stop translation

<400> 294

Met Met Tyr Val Gln Ser Ala Ile Met Ser Leu Gln His Leu Leu Val
1 5 10 15

Leu His Arg Val Ile Ile Ile Ser Met His Phe Ala Phe Gly Asn Gly
20 25 30

Cys Thr Phe Lys Ile Leu Val Gln Cys Ala Ile Arg Lys Tyr Thr Ser
35 40 45

Lys Met Ile Ser Arg Ile Ile Gln Met Tyr Leu Thr Thr Met Asp Leu

50

55

60

Phe His Pro Met Lys Leu Gln Arg Lys Leu Xaa
55 70 75

<210> 295
<211> 51
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (51)
<223> Xaa equals stop translation

<400> 295
Met Ile Ile Pro Lys Phe Tyr Leu Phe Lys Leu Leu Leu Leu Gln
1 5 10 15

Lys Ile Thr His Phe Ile Cys Gly Lys Thr Leu Asn Asn Leu Asn Phe
20 25 30

Arg Cys Glu Ser Tyr Phe Leu Phe Leu Tyr Leu Tyr Cys Ala Tyr Ile
35 40 45

Leu Tyr Xaa
50

<210> 296
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation

<400> 296
Met Thr Gln Glu Ile Leu Val Val Phe Ser Ile Gln Val Leu Ser Ser
1 5 10 15

Leu Arg Leu Leu Gly Leu Trp Phe Phe Met Glu Asn Arg Leu Cys Ser
20 25 30

Gly Ile Val Glu Gln Arg Arg Leu Leu His Leu Asn Xaa
35 40 45

<210> 297
<211> 48
<212> PRT
<213> Homo sapiens

<220>
<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 297

Met Pro Thr Leu Gly Asp Ala Leu Ile Leu Tyr Leu His Leu Val Leu
1 5 10 15

Gly Val Ala Gly Val Leu Gln Pro Pro Gly Pro Arg Pro Ser Gln Ala
20 25 30

Leu Gly Pro Thr Gly Asp Arg Ala Pro Gly Lys Trp Asn Arg Ser Xaa
35 40 45

<210> 298

<211> 55

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (55)

<223> Xaa equals stop translation

<400> 298

Met Ala Trp Cys Leu Leu Ser Val Phe Phe Leu Arg Ala Leu Cys Ala
1 5 10 15

His Ser Ser Thr Ala Tyr Lys Cys Val Leu Cys Ser Pro Arg Ser Pro
20 25 30

Trp Leu Val Glu Ala Asn Phe Trp Leu Asp Phe Tyr Gly Lys Ser Tyr
35 40 45

Phe Met Ser Pro Lys His Xaa
50 55

<210> 299

<211> 30

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (30)

<223> Xaa equals stop translation

<400> 299

Met Gln Met Thr Val Val Trp Tyr Val Ile Thr Ala Ile Ile Trp Trp
1 5 10 15

Arg Met Ser Met Cys Glu Ala Leu Ser Gln Asn Cys Phe Xaa
20 25 30

<210> 300

<211> 73

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (73)

<223> Xaa equals stop translation

<400> 300

Met Pro Leu Gly Val Val Pro Arg Ala Val Trp Ser Thr Leu Ala Trp
1 5 10 15

Val Cys Ile Ile Leu Gln Thr Leu Lys Thr Ser Leu Phe Cys Gln Thr
20 25 30

Thr Phe Cys Gly Glu Pro Glu Asp Ser Gly Phe Phe Glu Gly Ile Leu
35 40 45

Asp Val Cys Val Leu Val Lys Glu Ala Val Ile Arg Leu Asn His Asn
50 55 60

Pro Gln Asp Leu Leu Asp Ser Asp Xaa
65 70

<210> 301

<211> 37

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (37)

<223> Xaa equals stop translation

<400> 301

Met Leu Arg Leu Glu Val Leu Leu Phe Phe Ser Lys Val Thr Asp
1 5 10 15

Gln Ile Ile Thr Gln Ile Ile Gln Glu Asn Arg Ser Glu Ile Lys Asn
20 25 30

Asn Ile Ile Phe Xaa
35

<210> 302

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals stop translation

<400> 302

Met Arg Pro Val Leu Arg Arg Thr Phe Leu Leu Thr Leu Phe Ser Val
1 5 10 15

Ile Ala Leu Thr Lys Ile Lys His Asp Phe Phe Ile Met Cys Ser His
20 25 30

Met Gln Cys Ile Pro Arg Val Phe Leu Lys His Glu Phe Asn Asn Ile
35 40 45

Xaa

<210> 303

<211> 42

<212> PRT

<213> Homo sapiens

<400> 303

Met Phe Tyr Thr Thr Leu Cys Lys Met Phe Gln Tyr Leu His Ile Leu
1 5 10 15

Ser Leu Ser Phe Cys Phe Ala Leu Ile Trp Trp Ser Glu Ser Phe Leu
20 25 30

Trp Leu Ser Asn Leu Val Arg Leu Arg His
35 40

<210> 304

<211> 54

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (54)

<223> Xaa equals stop translation

<400> 304

Met Ile Leu Leu Ile Ser Gln Cys Pro Leu Ser Ile Phe Ala Ala Pro
1 5 10 15

Phe Ala Leu Pro Pro Lys Gly His Cys Gly Ser Phe Ser Asp Phe His
20 25 30

Ser Gln Val Thr Leu His Lys Asn Ser Lys Leu Ile Phe Arg Ser His
35 40 45

Lys Ser Ile Leu Leu Xaa
50

<210> 305

<211> 76

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (76)

<223> Xaa equals stop translation

<400> 305

Met Leu Ala Ala Glu Leu Ile Cys Cys Pro Ser Leu His Ile Phe Phe
1 5 10 15

Phe Ala Ala Phe Ser Leu Trp Gln Cys Thr Val Leu Thr Met Pro Phe
20 25 30

Lys Asn Val Pro Tyr Cys Ile Ser Ile Leu Arg Arg Asp Arg Thr Lys
35 40 45

Lys Tyr Ile Ala Gln Ile Ile Phe Tyr Phe Ile Asp Asn Asp Lys Glu
50 55 60

Tyr Phe Leu Asn Pro Ile Lys Ile Asp Phe Asn Xaa
65 70 75

<210> 306

<211> 63

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (63)

<223> Xaa equals stop translation

<400> 306

Met Phe Phe Arg Met Gln Val Cys Glu His His Gly Phe Trp Val Ile
1 5 10 15

Leu Leu Leu Leu Ser Leu Lys Met Glu Ile Pro Leu Ala Ala Tyr Pro
20 25 30

Thr Ala Glu Tyr Ser Ser Ile Gly Ser Gly Phe Thr Pro Leu His Pro
35 40 45

Ser Arg Thr Phe Thr Gln Ala Ser Pro Leu Pro Ser Ile Phe Xaa
50 55 60

<210> 307

<211> 50

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (50)

<223> Xaa equals stop translation

<400> 307

Met Asn Val Phe Val Gly Pro Leu Ser Val Ala Ile Val Ile Phe Cys
1 5 10 15

Trp Ile Thr Met Tyr Trp Val Ser Ile Val Met Gly Gln Gly Arg Gly
20 25 30

Gln Tyr Thr Trp Arg Thr Ile Leu Ser Thr Ser Thr Pro Ser Val Cys
35 40 45

Ser Xaa
50

<210> 308

<211> 103

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (103)

<223> Xaa equals stop translation

<400> 308

Met Glu His Trp Ile Pro Pro Glu Val Pro Leu Ala Gly Leu Arg Arg
1 5 10 15

Leu Leu Leu Asp Arg Leu Val Phe Ala Pro Ala Phe Leu Met Leu Phe
20 25 30

Phe Leu Ile Met Asn Phe Leu Glu Gly Lys Asp Ala Ser Ala Phe Ala
35 40 45

Ala Lys Met Arg Gly Gly Phe Trp Pro Ala Leu Arg Met Asn Trp Arg
50 55 60

Val Trp Thr Pro Leu Gln Phe Ile Asn Ile Asn Tyr Val Pro Leu Lys
65 70 75 80

Phe Arg Val Leu Phe Ala Asn Leu Ala Ala Leu Phe Trp Tyr Ala Tyr
85 90 95

Leu Ala Ser Leu Gly Lys Xaa
100

<210> 309

<211> 45

<212> PRT

<213> Homo sapiens

<400> 309

Met Arg Phe Ile Ser Gln Gln Ser Cys Glu Cys Val Arg Pro Cys Met
1 5 10 15

Asp Val Tyr Val Cys Val Tyr Ile Ser Ile His Val Tyr Met Asp Ala
20 25 30

His Val Tyr Leu Cys Arg Ile Cys Lys Thr Asn Met Arg
35 40 45

<210> 310

<211> 53

<212> PRT

<213> Homo sapiens

<400> 310

Arg Ile Leu Arg Trp Val Asn Cys Met Ala Cys Asp Leu Tyr Leu Asn
1 5 10 15

Lys Ala Val Ser Val Cys Ala His Val Trp Met Cys Met Cys Val Tyr
20 25 30

Ile Ser Leu Tyr Met Tyr Thr Trp Met Pro Met Cys Ile Tyr Val Glu
35 40 45

Tyr Val Lys Gln Thr

50

<210> 311

<211> 59

<212> PRT

<213> Homo sapiens

<400> 311

Asn Pro Glu Asn Gln Leu Glu Ile Ser Phe Pro Pro Arg Arg Gln Lys
1 5 10 15

Met Lys Leu Thr Leu Asp Leu Gln Val Ser Gln Ser Ser Leu Val His
20 25 30

Ser Leu Leu Ser Ser Asp Phe Phe Ser Val Ser Lys Glu Gly Cys Leu
35 40 45

Trp Lys Pro Ile Leu Leu Pro Ser His Phe Leu
50 55

<210> 312

<211> 47

<212> PRT

<213> Homo sapiens

<400> 312

Leu Gln Thr Gln Ile Ser Asn Tyr Leu Met Phe Val Leu His Ile Leu
1 5 10 15

His Arg Tyr Thr Trp Ala Ser Met Tyr Thr Cys Ile Glu Ile Tyr Thr
20 25 30

His Thr Tyr Thr Ser Ile His Gly Arg Thr His Ser Gln Leu Cys
35 40 45

<210> 313

<211> 45

<212> PRT

<213> Homo sapiens

<400> 313

Ile His Met Gly Ile His Val Tyr Met Tyr Arg Asp Ile Tyr Thr His
1 5 10 15

Ile His Ile His Thr Trp Ala His Thr Leu Thr Ala Leu Leu Arg Tyr
20 25 30

Lys Ser His Ala Ile Gln Leu Thr His Leu Asn Ile Arg
35 40 45

<210> 314

<211> 41

<212> PRT

<213> Homo sapiens

<400> 314

Met Lys Trp Ile Phe Thr Val Leu Ile Leu Thr Ser Cys Phe Phe Thr
1 5 10 15

Ala Gly Ile Cys Glu Asp Gly Ile Cys Ser Arg Ile Gln Leu Arg Asp
20 25 30

Lys Ile Val Gln Ser Ala Phe Arg Gln
35 40

<210> 315

<211> 81

<212> PRT

<213> Homo sapiens

<400> 315

Lys Pro Cys Cys Pro Ser Val Ser Asn Arg Ser Ser Val Gln Met His
1 5 10 15

Gln Leu Pro Ile Gln Phe Leu Gly Gln Phe Glu Ala His Cys Ile Gly
20 25 30

Phe Cys Arg Ser Phe Leu Glu Thr Phe Tyr Thr His Asp Pro Arg Ala
35 40 45

Met His Ser Phe Leu Ser Ser Ile Ser Ser Pro Ser Leu Pro Phe Gly
50 55 60

Phe Ser Arg Met Thr Ser Gln Ile Asn His Leu His Pro Ser Pro Leu
65 70 75 80

Cys

<210> 316
<211> 21
<212> PRT
<213> Homo sapiens

<400> 316
Ser Val Phe Lys Ile Asn Leu Lys Ser Phe Lys Gln His Glu Pro Trp
1 5 10 15
Trp Pro Asn Arg Ser
20

<210> 317
<211> 135
<212> PRT
<213> Homo sapiens

<400> 317
Gly Thr Arg Ser Phe Ser Val Pro Ser Tyr Leu Arg Leu Thr Gly Ser
1 5 10 15

Leu Met Cys Tyr Leu Leu Leu Leu Ile Gln Thr Ala Glu Leu Leu
20 25 30

Ile His Pro Gln Gly Leu Gln Ala Val Ser Asn Gly Glu Ser Ala Leu
35 40 45

Lys Gly Thr Arg Pro Thr Phe Ser Ser Pro Phe Ile Leu Val Thr Glu
50 55 60

Gly Arg Lys Glu Trp Glu Gly Val Phe Leu Ser Ser Gly Trp Lys Gly
65 70 75 80

Asn Thr Leu Ser Asn Tyr Tyr Ile Ser Leu Val Phe Tyr Tyr Ser Arg
85 90 95

Ile Leu Gln Pro Tyr Phe Tyr Cys Leu Trp Gly Lys Leu Glu Met Val
100 105 110

Thr Leu Ile Arg Ser Val Trp Arg Gly Ile Asn Gly Gly Asp Lys Ile
115 120 125

Ser Val Gly Phe Gly Lys Cys
130 135

<210> 318
<211> 38
<212> PRT
<213> Homo sapiens

<400> 318
Trp Met Glu Arg Lys His Thr Val Lys Leu Leu Tyr Leu Leu Gly Phe
1 5 10 15

Leu Leu Gln Asn Ser Pro Ala Ile Phe Leu Leu Ser Met Gly Glu Val
20 25 30

Gly Asp Gly Asp Leu Asp
35

<210> 319

<211> 23

<212> PRT

<213> Homo sapiens

<400> 319

Ser Asn Gly Glu Ser Ala Leu Lys Gly Thr Arg Pro Thr Phe Ser Ser
1 5 10 15

Pro Phe Ile Leu Val Thr Glu
20

<210> 320

<211> 24

<212> PRT

<213> Homo sapiens

<400> 320

Leu Ser Asn Tyr Tyr Ile Ser Leu Val Phe Tyr Tyr Ser Arg Ile Leu
1 5 10 15

Gln Pro Tyr Phe Tyr Cys Leu Trp
20

<210> 321

<211> 131

<212> PRT

<213> Homo sapiens

<400> 321

Glu Lys Asp Phe Met Gln Gly Ser Asp Ala Gly His Gly Gly Thr His
1 5 10 15

Ile Tyr Arg Ala Leu Val Gln Trp Pro Leu Ala Trp Val Phe Tyr Leu
20 25 30

Ser His Ala Lys Thr His Trp Gly Glu Glu Leu Arg Phe Ser Phe Arg
35 40 45

Arg Lys Asn Leu Arg Leu Arg Glu Ala Met Arg His Glu Thr Cys Gln
50 55 60

Val Thr Gln Leu Val Ala Gly Lys Ala Asp Ser Asn Leu Cys Leu Arg
65 70 75 80

Asp Ser Glu Thr Trp Phe Trp Pro Pro Leu Trp Ala Ala Cys Ser Ser
85 90 95

Leu Gln Ala Thr Ala Cys Arg Leu Ser Ser Pro Ser Lys Gly Leu Gly
100 105 110

Ala Ser Arg Glu Cys Pro Trp Leu Ala Ser Gly Arg Ala Ala Leu Val
115 120 125

Ser Phe Leu
130

<210> 322
<211> 69
<212> PRT
<213> Homo sapiens

<400> 322
Ser Leu Arg Val Lys Gly Arg Lys Pro Arg Leu Leu Tyr His Ser Pro
1 5 10 15

Ala Arg Gly Thr Leu Trp Met Leu Pro Gly Leu Cys Asp Cys Leu Ile
20 25 30

Cys Arg Gln Trp Leu Val Glu Arg Ser Arg Leu Pro Arg Val Gly Ala
35 40 45

Arg Thr Arg Phe Gln Ser Pro Ser Asp Thr Gly Trp Ser Gln Leu Cys
50 55 60

Gln Leu Pro Ala Val
65

<210> 323
<211> 26
<212> PRT
<213> Homo sapiens

<400> 323
Glu Arg Ser Arg Leu Pro Arg Val Gly Ala Arg Thr Arg Phe Gln Ser
1 5 10 15

Pro Ser Asp Thr Gly Trp Ser Gln Leu Cys
20 25

<210> 324
<211> 33
<212> PRT
<213> Homo sapiens

<400> 324
Lys His Ala Phe Leu Met Ala His Gln Phe Cys Val Leu Ser Leu Ala
1 5 10 15

Met Gln Trp Ser Ser Cys Phe Gln Leu Val Ala Leu Pro Tyr Leu Ser
20 25 30
Leu

<210> 325
<211> 51
<212> PRT
<213> Homo sapiens

<400> 325
Met Arg Pro Leu Cys Val Leu Leu Pro Trp Pro Cys Trp Gln Trp Gly
1 5 10 15

Gly Leu Gly Ser Ala Ser Pro Ile Arg Pro Gln Ala Pro Pro Gly Gln
20 25 30

Ala Ala His Ala Val Pro Leu Pro Arg Ala Gln His Leu Ala Gln Arg
35 40 45

Ser Arg Gln
50

<210> 326
<211> 52
<212> PRT
<213> Homo sapiens

<400> 326
Ala Arg Gly Leu Arg Ser Pro His Gly Ala Ala Gly Val Val Arg Gly
1 5 10 15

Asp Gly Gly Gly Lys Lys Gly Glu Asp Pro Tyr Ser Pro Ile Leu Phe
20 25 30

Gln Ser Glu Arg Ile Pro Arg Leu Ile Tyr Leu Pro Val Ile Ser Ser
35 40 45

Glu Glu Asn Ser
50

<210> 327
<211> 57
<212> PRT
<213> Homo sapiens

<400> 327
Lys Ser Leu Ser Cys Ser Phe Leu Phe Leu Ala Phe Trp Leu Arg Arg
1 5 10 15

Met Gly Gin Thr Met Cys Val Cys Val Cys Val Cys Val Cys Val Cys
20 25 30

Val Arg Thr Trp Val Tyr Leu Tyr Glu Pro Val Lys Phe Arg Ser Pro
35 40 45

Leu Ile Tyr Val Asn Leu Pro Thr Ser
50 55

<210> 328

<211> 80
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (15)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 328
Lys Leu Gly Phe Thr Met Leu Ala Arg Leu Val Ser Asn Ser Xaa Thr
1 5 10 15

Ser Gly Asp Leu Pro Ser Ser Ala Ser Gln Asn Ala Gly Ile Lys Gly
20 25 30

Met Ser Tyr Arg Ala Trp Pro Tyr Ser Tyr Phe Leu Ile Arg Lys Asn
35 40 45

Lys Gln Thr Asn Lys Gln Thr Lys Thr Asn Pro Gln Leu Gly Glu Asn
50 55 60

Lys His Cys Arg Asn Leu Lys Val Ser Trp Ser Lys Asn Tyr Phe Leu
65 70 75 80

<210> 329
<211> 27
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (25)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 329
Glu Arg Gly Gln Gly Gly Ser Ser Arg Asn Val Ala Gly Ser Asp Leu
1 5 10 15

Val Phe Pro Ala Val Phe Val Ser Xaa Leu Cys
20 25

<210> 330
<211> 166
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (90)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE
<222> (92)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (96)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (113)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (126)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (141)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (150)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 330
Gly Ser Pro Gln Gly Pro Ser Val Ala Leu Gly Ser Arg Gln Cys Trp
1 5 10 15

Ser Arg Pro Leu Arg Arg Gly Gly Arg Gly Ala Ala Val Glu Met Trp
20 25 30

Arg Gly Pro Thr Trp Cys Phe Arg Pro Ser Leu Cys Leu Cys Cys Val
35 40 45

Cys Gly Val Ser Phe Gly Leu Tyr Val Pro His Gly Phe Ser Leu Ser
50 55 60

Met Cys Val Ser Ala Pro Gly Ser Ala Trp Leu Ser Leu Val Tyr Ser
65 70 75 80

Ile Cys Leu Ala Arg Gly Ser Met Ser Xaa Arg Xaa Ser Ser Arg Xaa
85 90 95

Ser Leu Val Ala Ser Gly Ala Ser Val Leu Leu Val Cys Phe Trp Val
100 105 110

Xaa Ala Asp Pro Gly Val Gly Val Ser Val Pro Arg Ala Xaa Val Ser
115 120 125

Gly Leu Trp Trp Cys Val Ser Pro Ser Ala Cys Leu Xaa Leu Ala Pro
130 135 140

Thr Lys Pro Pro Pro Xaa Leu Ser Phe Ser Leu Ser Ile Phe Pro Phe

145

150

155

160

Ser Ser Asn Pro Ser Lys
165

<210> 331
<211> 118
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (31)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (39)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (55)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (67)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (84)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (89)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (90)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 331
Thr Ile Ala Ser Leu Gln Pro Thr Ala Leu Asn His Leu Ile Trp Arg
1 5 10 15

Gly Trp Lys Arg Lys Gly Arg Leu Arg Glu Arg Lys Arg Gly Xaa Gly
20 25 30

Gly Ala Trp Leu Gly Pro Xaa Arg Gly Arg Gln Met Asp Ser His Thr
35 40 45

Thr Arg Asp Gln Arg Gln Xaa Leu Gly Glu Gln Arg His Pro Leu Leu
50 55 60

Gly Leu Xaa Ala Pro Arg Ser Lys Pro Thr Lys Gln Met Pro Gln Met
65 70 75 80

Gln Pro Gly Xaa Pro Glu Lys Lys Xaa Xaa Leu Thr Trp Asn His Gly
85 90 95

Leu Asp Arg Trp Asn Thr Gln Gly Thr Ala Arg Gln Ser Leu Gly Gln
100 105 110

Lys His Thr Trp Arg Asp
115

<210> 332

<211> 21

<212> PRT

<213> Homo sapiens

<400> 332

Ala Arg Gly Pro Gly Thr Glu Gly Cys Glu Pro Trp Leu Gln Leu Gln
1 5 10 15

Asp Arg Arg Glu Arg
20

<210> 333

<211> 59

<212> PRT

<213> Homo sapiens

<400> 333

Met Ser Ser Gly Thr Asn Ser Phe Phe Thr Leu Met Ala Leu Asn Ser
1 5 10 15

Pro Thr Gly Asp Ser Gly Ser Arg Ile Thr Val Ser Pro Pro Arg Val
20 25 30

His Pro Val Lys Ser Gly Arg Gly Arg Ala Ser Asp Leu Leu Leu Thr
35 40 45

Arg Phe Leu Ala Pro Arg Ser Ala Leu Trp Ser
50 55

<210> 334

<211> 26

<212> PRT

<213> Homo sapiens

<400> 334

His Glu Tyr His Leu Leu Ser Ser Arg His Ile Leu Gly Ser Val Leu
1 5 10 15

Arg Leu Asp Val Cys Ser Ala Leu Trp Ser
20 25

<210> 335
<211> 82
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (44)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (54)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (59)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (67)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 335
Phe Ile Leu Phe Ile Leu Glu Tyr Asp Met Leu Trp Lys Ser Leu Tyr
1 5 10 15

Thr Asn Ser Ser Ala Tyr Gly Tyr Val Ile Ala Ser Tyr Phe Cys Leu
20 25 30

Leu Gly Ile Lys Leu Leu Val Lys Gln Lys Lys Xaa Lys Lys Lys Thr
35 40 45

Arg Gly Gly Ala Arg Xaa Pro Ile Arg Pro Xaa Val Glu Ser Tyr Tyr
50 55 60

Lys Ser Xaa Ala Val Val Leu Gln Arg Arg Gly Leu Gly Lys Asn Leu
65 70 75 80

Gly Gly

<210> 336
<211> 102
<212> PRT
<213> Homo sapiens

<400> 336
Arg Val Ser Ser His Leu Phe Arg Leu Phe Gly Gly Leu Ile Leu Asp
1 5 10 15

Ile Lys Arg Lys Ala Pro Phe Phe Leu Ser Asp Phe Lys Asp Ala Leu
20 25 30

Ser Leu Gln Cys Leu Ala Ser Ile Leu Phe Leu Tyr Cys Ala Cys Met
35 40 45

Ser Pro Val Ile Thr Phe Gly Gly Leu Leu Gly Glu Ala Thr Glu Gly
50 55 60

Arg Ile Val Ser Thr Lys Ile Gly Ser Gly Gln Ala Phe Ser Ser Ser
65 70 75 80

Glu Ala Ser Val Cys Met His Leu Ser His Tyr Ser Tyr Phe Tyr Leu
85 90 95

Lys Ser Leu Pro Thr Ala
100

<210> 337

<211> 24

<212> PRT

<213> Homo sapiens

<400> 337

Phe Arg Leu Phe Gly Gly Leu Ile Leu Asp Ile Lys Arg Lys Ala Pro
1 5 10 15

Phe Phe Leu Ser Asp Phe Lys Asp
20

<210> 338

<211> 23

<212> PRT

<213> Homo sapiens

<400> 338

Phe Leu Tyr Cys Ala Cys Met Ser Pro Val Ile Thr Phe Gly Leu
1 5 10 15

Leu Gly Glu Ala Thr Glu Gly
20

<210> 339

<211> 22

<212> PRT

<213> Homo sapiens

<400> 339

Ser Ser Ser Glu Ala Ser Val Cys Met His Leu Ser His Tyr Ser Tyr
1 5 10 15

Phe Tyr Leu Lys Ser Leu
20

<210> 340

<211> 106

<212> PRT

<213> Homo sapiens

<400> 340

Pro Cys Leu Gln Val Ile Gly Ile Asp Phe Cys Arg Leu Leu Met
1 5 10 15

Cys Leu Val Leu Lys Arg Asn Leu Thr Val Pro Phe Ser Ser Tyr Ser
20 25 30

Pro Leu Lys Thr Ile Thr Cys Ile Thr Ser Glu Gln Ile Ala Val Val
35 40 45

Ser Asn Phe Phe Arg Gln Lys Leu Gly Val Arg Ala Lys Phe Phe Gln
50 55 60

Gly Ala Cys Leu His Thr Ser Lys Val Val Ile Cys Leu Asn Leu Pro
65 70 75 80

Ile Ile Ser Ile Gln Arg Ala Asp Ile Arg Met Trp Trp Leu Val Val
85 90 95

Asn Thr Pro Tyr Ala Arg Gly Val Asn Asn
100 105

<210> 341

<211> 21

<212> PRT

<213> Homo sapiens

<400> 341

Val Val Ser Val Cys Val Leu Glu Thr Gly Gln Leu Gly Pro Ala Ala
1 5 10 15

Leu Cys Arg Ser Val
20

<210> 342

<211> 97

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (79)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (83)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (85)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (90)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 342

Asn	Ile	Ser	Val	His	Gly	Phe	Pro	Val	Pro	Cys	Leu	Arg	Gln	Arg	Leu
1									10					15	

Gln	Gly	Pro	Cys	His	Pro	Lys	Cys	Cys	Pro	His	Xaa	Ile	Ser	Ser	Gly
									25					30	

Lys	Pro	Arg	Ser	Ser	Phe	Ser	Pro	Ser	Ser	Tyr	His	Cys	Lys	Phe	Ser
									40					45	

Arg	Asn	Ala	Thr	Leu	Leu	Val	Val	Pro	Asn	Ile	Phe	Ser	Tyr	Met	Gln
								55					60		

Ser	Ser	Phe	Leu	Ile	Pro	Gln	Thr	Ser	Lys	Tyr	Tyr	Ile	Leu	Xaa	Pro
65								70				75			80

Tyr	Ala	Xaa	Thr	Xaa	Arg	Pro	Ile	Lys	Xaa	Ile	Phe	Lys	Gln	Ala	Lys
								85				90			95

Gln

<210> 343

<211> 58

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (19)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 343

Ile	Tyr	Asn	Asp	Met	Met	Glu	Lys	Lys	Thr	Glu	Val	Tyr	Gln	
1					5				10				15	

Lys	Arg	Xaa	Ser	Gly	Asp	Asn	Thr	Trp	Gly	Gly	Lys	Gly	Leu	Val	Ala
								20			25			30	

Phe	Val	Ser	Ser	Met	Glu	Gln	Gly	Ile	His	Val	Gln	Arg	Cys	Phe	Ile
								35		40			45		

Ala	Asn	Leu	Lys	Phe	Ser	Ser	Pro	Gly	Val					
									50				55	

<210> 344

<211> 93
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (16)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 344
Tyr Asp Asp Gly Glu Lys Glu Asp Arg Gly Leu Pro Glu Glu Met Xaa
1 5 10 15

Trp Gly Gln His Leu Gly Trp Gln Gly Pro Cys Ser Leu Cys Leu Lys
20 25 30

His Gly Thr Gly Asn Pro Cys Thr Glu Met Phe Tyr Cys Gln Phe Lys
35 40 45

Ile Phe Ile Ser Trp Cys Leu Ile Pro Leu Val Phe Ala Arg Leu Gly
50 55 60

Asp Phe Arg Asp Arg Pro Gly Trp Ile Phe Ser Trp Arg Tyr His Leu
65 70 75 80

Lys His Thr Val Trp Gly Gly Tyr Asn Ile Ile Met Leu
85 90

<210> 345
<211> 21
<212> PRT
<213> Homo sapiens

<400> 345
Thr Pro Gly Asp Glu Asn Phe Lys Leu Ala Ile Lys His Leu Cys Thr
1 5 10 15

Trp Ile Pro Cys Ser
20

<210> 346
<211> 34
<212> PRT
<213> Homo sapiens

<400> 346
Ile Arg His Glu Ile Phe Leu Thr Ile Glu Ser Phe Cys Pro Ser Ala
1 5 10 15

Pro Arg Gly Glu Asp Asp Asp Asn Leu Leu Arg Thr Ser Arg Val Pro
20 25 30

Asp Ile

<210> 347
<211> 160
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (126)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (130)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 347
Ile Arg Gly Ser Ile Pro Gly His Lys Lys Met His Leu Ser Phe Asn
1 5 10 15

Val Ala Ala Gln Trp Ser Leu Leu Lys Pro Leu Val Leu Arg Glu Glu
20 25 30

Gly Ala Leu Phe Leu Thr His Asp Gln Leu Glu Ser Lys Asn Ser Trp
35 40 45

Thr Leu Ser Ile Gly Pro Arg Val Pro Tyr Thr Tyr Val Val Val Thr
50 55 60

Trp Ser Ser Ala Leu Trp Asp Leu Pro Asn Gln Pro Leu Ala Gly Arg
65 70 75 80

Lys Glu Ser Gly Gly Ser Tyr Gly Pro Ile Ser Val Thr Gln Ser Pro
85 90 95

His Gln Ala Ala Leu Lys Trp Phe Ala Lys Lys Gly Lys Gln Ser
100 105 110

His Ser Thr Val Gln Leu Ala Asn Ile Leu His Val Phe Xaa Ala Pro
115 120 125

Asp Xaa Tyr His Phe Val Asn Thr Ser Leu Gln Leu Phe Leu Glu Tyr
130 135 140

Thr Val Met Cys Met Leu Cys Glu Asn Lys Gln Lys Thr Leu Gly Arg
145 150 155 160

<210> 348
<211> 135
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (8)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (10)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 348

Glu Pro Glu Val Thr Gln Val Xaa Ser Xaa Glu Leu Thr Phe Gln Pro
1 5 10 15

Arg Lys Ala Gly Ala Lys Val Thr Ala Gly Lys Ser His His Gln Val
20 25 30

Ile His Trp Glu Phe Glu Ile Met Leu Ser Ser Tyr Ser Thr Asp Val
35 40 45

Pro Leu Trp Phe Leu Lys Phe Phe Ser Ser Asn Leu Pro Gln Thr Tyr
50 55 60

Phe Pro His Ser Gly Val Lys Lys Trp Gly Ser Cys Phe Ser Leu Pro
65 70 75 80

Trp Arg Asp Ser Pro Pro Leu Thr Phe Ile Ser Leu Leu Ser Ser His
85 90 95

Leu Thr Thr Phe His Leu Tyr His Leu His His Gly Ile Ile Cys Leu
100 105 110

Gly Phe Ser Val Tyr Phe His Arg Ala Tyr Thr Ser Leu Cys Ile Leu
115 120 125

Glu Thr Ala Val Gly Ser Tyr
130 135

<210> 349

<211> 25

<212> PRT

<213> Homo sapiens

<400> 349

Trp Ser Leu Leu Lys Pro Leu Val Leu Arg Glu Glu Gly Ala Leu Phe
1 5 10 15

Leu Thr His Asp Gln Leu Glu Ser Lys
20 25

<210> 350

<211> 22

<212> PRT

<213> Homo sapiens

<400> 350

Trp Phe Ala Lys Lys Lys Gly Lys Gln Ser His Ser Thr Val Gln Leu
1 5 10 15

Ala Asn Ile Leu His Val
20

<210> 351
<211> 25
<212> PRT
<213> Homo sapiens

<400> 351
Ala Gly Lys Ser His His Gln Val Ile His Trp Glu Phe Glu Ile Met
1 5 10 15

Leu Ser Ser Tyr Ser Thr Asp Val Pro
20 25

<210> 352
<211> 26
<212> PRT
<213> Homo sapiens

<400> 352
His Gly Ile Ile Cys Leu Gly Phe Ser Val Tyr Phe His Arg Ala Tyr
1 5 10 15

Thr Ser Leu Cys Ile Leu Glu Thr Ala Val
20 25

<210> 353
<211> 19
<212> PRT
<213> Homo sapiens

<400> 353
Lys Arg Leu Thr Ile Asn Ala Arg Val His Leu Trp Thr Leu Lys Ser
1 5 10 15
Val Pro Leu

<210> 354
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (7)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (8)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 354

Glu Tyr Val Phe Asn Met Xaa Xaa Tyr Ser Lys Ser Arg Ala Ile Ser
1 5 10 15

Pro Leu Ser Gly Pro Tyr Thr Pro Arg Gly Thr Thr Pro Leu Pro Ile
20 25 30

Ile Pro Glu Pro Gly Ala Arg Gln Arg Asp His Pro Ala Ser Leu Lys
35 40 45

Tyr Ala Lys Ile Ile Gln Thr Lys Leu Phe Ala Leu Pro Tyr Pro Lys
50 55 60

Glu Thr Ser Met Lys Ala Val Ala
65 70

<210> 355

<211> 65

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (15)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (25)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (26)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 355

Glu Thr Val Pro Pro Arg Ser Ser Gln Phe Leu Lys Ile Thr Xaa Gly
1 5 10 15

Pro Ala Arg Ser Met Ser Leu Ile Xaa Xaa Ala Ile Gln Asn Pro Glu
20 25 30

Pro Tyr Leu Leu Tyr Leu Ala Leu Ile Pro Gln Glu Ala Leu Leu Leu
35 40 45

Tyr Leu Ser Ser Gln Ser Gln Val Pro Gly Asn Glu Thr Thr Pro Pro
50 55 60

Val

65

<210> 356

<211> 101

<212> PRT

<213> Homo sapiens

<400> 356

Asn	Glu	Val	Ser	Phe	Ser	Leu	Ser	Leu	Gly	Phe	Ser	Pro	Arg	Glu	Phe
1				5					10					15	

Ala	Arg	Trp	Lys	Val	Asn	Asn	Leu	Ala	Leu	Glu	Arg	Lys	Asp	Phe	Phe
				20					25				30		

Ser	Leu	Pro	Leu	Pro	Leu	Ala	Pro	Glu	Phe	Ile	Arg	Asn	Ile	Arg	Leu
							35				40		45		

Leu	Gly	Arg	Arg	Pro	Asn	Leu	Gln	Gln	Val	Thr	Glu	Asn	Leu	Ile	Lys
					50			55			60				

Lys	Tyr	Gly	Thr	His	Phe	Leu	Leu	Ser	Ala	Thr	Leu	Gly	Gly	Lys	Gln
				65			70			75				80	

His	His	Asn	Pro	Lys	Leu	Ile	Gly	Cys	Gln	Thr	Ile	Gly	Asn	Asn	Val
					85				90		95				

Lys	Thr	Arg	Val	Ala											
				100											

<210> 357

<211> 75

<212> PRT

<213> Homo sapiens

<400> 357

Val	Pro	Tyr	Phe	Leu	Ile	Arg	Phe	Ser	Val	Thr	Cys	Cys	Arg	Leu	Gly
1				5					10				15		

Leu	Leu	Pro	Arg	Arg	Arg	Met	Phe	Arg	Ile	Asn	Ser	Gly	Ala	Arg	Gly
						20			25			30			

Asn	Gly	Lys	Leu	Lys	Lys	Ser	Phe	Leu	Ser	Arg	Ala	Lys	Leu	Phe	Thr
						35			40			45			

Phe	Gln	Arg	Ala	Asn	Ser	Leu	Gly	Glu	Lys	Pro	Arg	Asp	Lys	Glu	Lys
				50				55			60				

Leu	Thr	Ser	Phe	Gln	Ser	Lys	Arg	His	Lys	Ile					
					65			70		75					

<210> 358

<211> 63

<212> PRT

<213> Homo sapiens

<400> 358

Glu	Met	Ser	Ala	Val	Leu	Phe	Asn	Gln	Ile	Phe	Cys	Asn	Leu	Gln	
1				5					10			15			

Ile	Gly	Ser	Pro	Ser	Lys	Glu	Ala	Asn	Val	Pro	Asp	Lys	Leu	Trp	Gly
					20				25			30			

Lys Arg Gln Trp Gln Thr Glu Glu Val Leu Pro Phe Gln Ser Gin Val
 35 40 45

Val His Leu Pro Thr Gly Lys Leu Pro Gly Gly Lys Ala Lys Gly
 50 55 60

<210> 359

<211> 99

<212> PRT

<213> Homo sapiens

<400> 359

His Tyr His Gly Ser Gly Phe Leu Ile Lys Glu Phe Gly Ser Phe Leu
 1 5 10 15

Ser Leu Leu Cys Met Leu Ser Cys Pro Tyr Val Phe Cys His Gly Met
 20 25 30

Leu Glu Gln Glu Val Pro Ser Ser Val Val Ser Pro Ser Thr Leu Asp
 35 40 45

Phe Pro Thr Ser Arg Thr Val Asn Lys Phe Leu Phe Lys Leu Pro Ser
 50 55 60

Leu Trp Tyr Ser Val Ile Ala Thr Gln Asn Gly Leu Lys Gln Lys Ile
 65 70 75 80

Arg Glu Thr Phe Leu Phe Val Gln Phe Ser Gln Met Pro Arg Trp His
 85 90 95

Lys Leu Glu

<210> 360

<211> 48

<212> PRT

<213> Homo sapiens

<400> 360

Phe Cys Lys His Asn Gly Ser Lys Asn Val Phe Ser Thr Phe Arg Thr
 1 5 10 15

Pro Ala Val Leu Phe Thr Gly Ile Val Ala Leu Tyr Ile Ala Ser Gly
 20 25 30

Leu Thr Gly Phe Ile Gly Leu Glu Val Val Ala Gln Leu Phe Asn Cys
 35 40 45

<210> 361

<211> 139

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (115)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 361

Met	Pro	Lys	Pro	Gly	Ala	Ala	Thr	Gln	Arg	Thr	Leu	Leu	Cys	Leu	Pro
1					5				10				15		

Arg	Leu	His	Pro	Ala	Ser	Gly	Pro	Pro	Leu	Pro	Xaa	Ala	Gly	Pro	Leu
					20				25				30		

Arg	Gly	Leu	Arg	Gln	Leu	Pro	Ala	Leu	Pro	Val	Pro	Ala	Ala	Ser	Cys
					35			40				45			

Arg	Arg	Arg	Pro	Ala	Pro	Arg	Leu	Cys	Ala	Ala	Gly	Pro	Cys	Thr	Val
					50		55				60				

Gly	Pro	Ala	Ala	Ser	Pro	His	Ala	Pro	Pro	His	Gly	Cys	Pro	Pro	Pro
					65		70			75			80		

Ala	Ser	Leu	Ala	His	Val	Ala	His	Arg	Gln	Ser	Val	Ser	Gly	Thr	Val
					85			90				95			

Cys	Leu	Gly	Leu	Arg	Asp	Gly	His	Val	Arg	Gly	Gly	Cys	Ala	Ala	Val
					100			105				110			

Arg	Gly	Xaa	Ala	Ala	Leu	Pro	Trp	Asp	Ala	Ala	Ala	Gly	Pro	Asp	
					115		120			125					

His	Met	Gly	Val	Gly	Ser	Gly	Pro	Ala	Leu	Leu					
					130		135								

<210> 362

<211> 35

<212> PRT

<213> Homo sapiens

<400> 362

Met	Trp	Gly	Gln	Pro	Arg	Pro	Val	Asp	Ser	Val	Trp	Ser	Ser	Ser	Ile
1						5			10			15			

Pro	Lys	Lys	Ser	Val	Glu	Ser	Asn	Asp	Asn	Lys	Ser	His	Leu	His	Lys
					20			25				30			

Arg	Glu	His													
		35													

<210> 363

<211> 26

<212> PRT
<213> Homo sapiens

<400> 363
Met Thr Thr Lys Ala Ile Phe Thr Lys Gly Asn Ile Asp Ser Leu Ser
1 5 10 15

Phe Lys Ser Asn Met Trp Ser Val Tyr Ile
20 25

<210> 364
<211> 26
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (3)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 364
Asp Ser Xaa Leu Asp Arg Arg Pro Ser Gly Pro Asp Val Lys Phe Leu
1 5 10 15

Ser Asn Lys His His Phe Ser Met Val Cys
20 25

<210> 365
<211> 84
<212> PRT
<213> Homo sapiens

<400> 365
Cys Leu Ala Glu Ala Val Ser Val Ile Gln Ser Ile Pro Ile Phe Asn
1 5 10 15

Glu Thr Gly Arg Phe Ser Phe Thr Leu Pro Tyr Pro Val Lys Ile Lys
20 25 30

Val Arg The Ser Phe Phe Leu Gln Ile Tyr Leu Ile Met Ile Phe Leu
35 40 45

Gly Leu Tyr Ile Asn Phe Arg His Leu Tyr Lys Gln Arg Arg Arg Arg
50 55 60

Tyr Gly Gln Lys Lys Lys Arg Ser Thr Lys Lys Lys Asp Leu Asp Gly
65 70 75 80

Phe Leu Pro Val

<210> 366
<211> 62
<212> PRT
<213> Homo sapiens

<400> 366

Leu Cys Ser Thr Pro Val Pro Thr Leu Phe Cys Pro Arg Ile Val Leu
1 5 10 15

Glu Val Leu Val Val Leu Arg Ser Ile Ser Glu Gln Cys Arg Arg Val
20 25 30

Ser Ser Gln Val Thr Val Ala Ser Glu Leu Arg His Arg Gln Trp Val
35 40 45

Glu Arg Thr Leu Arg Ser Arg Gln Arg Gln Asn Tyr Leu Arg
50 55 60

<210> 367

<211> 48

<212> PRT

<213> Homo sapiens

<400> 367

Ala Arg Gly Glu Thr Ala Tyr Asp Gly Ala Ala Val Glu Phe Gln Glu
1 5 10 15

Pro Leu Ser Ser Cys Leu Phe Ser Ser Leu Asn Pro His His Trp Pro
20 25 30

Thr Leu Gly Val Gly Arg Pro Val Met Leu Thr Leu Glu Asp Lys Asp
35 40 45

<210> 368

<211> 200

<212> PRT

<213> Homo sapiens

<400> 368

Glu Leu Leu Gln Cys Gln Met Leu Glu Ala Ser Thr Leu Ile His Leu
1 5 10 15

His His Pro Arg Pro Gly Phe Pro Ala Leu Cys Ser Phe Leu Gly Phe
20 25 30

Arg His His Leu His His Asp Ala Leu Cys Ile Arg Val Leu Pro Glu
35 40 45

Asp Leu Glu Ala Lys Leu Cys Val Ser Leu His Gln Leu Leu His Arg
50 55 60

Gly Leu Cys Leu Pro Gly Phe Gly Ala Ala Cys Pro Gly Asp Gln Gly
65 70 75 80

Ser Glu Asp Glu Ala Arg Pro Pro Ala Val Leu Arg Ala Val Ala Leu
85 90 95

Leu Arg Ala Gly Leu Arg His Leu Ser Val His Ser Gly Trp Tyr His
 100 105 110

Leu Pro His Ser Arg Asn Gly Leu Pro Leu Leu Ala Leu Val Val His
 115 120 125

Phe Pro Glu Tyr Gly Gly Pro Arg Glu Pro Val Pro Gly Gln Ser
 130 135 140

Gly Glu Phe Gly Arg Arg Thr Glu Leu Ser Thr Lys Gly Asp Thr Gly
 145 150 155 160

Asp Ser Arg Asn Ser His Leu Ala Gln Asp Met Ala Ser Leu Pro Phe
 165 170 175

Phe Lys Pro Cys Glu Cys Thr His Val Ala Val Cys Ser Pro Pro His
 180 185 190

Pro Leu Cys Gln Tyr Leu Cys Leu
 195 200

<210> 369

<211> 28

<212> PRT

<213> Homo sapiens

<400> 369

Leu Gln Cys Gln Met Leu Glu Ala Ser Thr Leu Ile His Leu His His
 1 5 10 15

Pro Arg Pro Gly Phe Pro Ala Leu Cys Ser Phe Leu
 20 25

<210> 370

<211> 31

<212> PRT

<213> Homo sapiens

<400> 370

His Gln Leu Leu His Arg Gly Leu Cys Leu Pro Gly Phe Gly Ala Ala
 1 5 10 15

Cys Pro Gly Asp Gln Gly Ser Glu Asp Glu Ala Arg Pro Pro Ala
 20 25 30

<210> 371

<211> 27

<212> PRT

<213> Homo sapiens

<400> 371

Leu Ala Leu Val Val His Phe Pro Glu Tyr Gly Gly Pro Arg Glu
 1 5 10 15

Pro Val Pro Gly Gln Ser Gly Glu Phe Gly Arg

20

25

<210> 372

<211> 30

<212> PRT

<213> Homo sapiens

<400> 372

Gln Ser Trp Thr Ala Pro Ala Ala Arg Leu Pro Met Ala Leu Pro Gln
1 5 10 15

Met Cys Asp Gly Ser His Leu Ala Ser Thr Leu Arg Tyr Cys
20 25 30

<210> 373

<211> 190

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (32)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (47)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 373

Gln Ser Ala Ala Gln Trp Phe Trp Trp Pro Gly Arg Ser Ala Ser Leu
1 5 10 15

Gly Gly Ala Lys Gly Met Gln Pro Pro Ser Leu Ala Ser Trp Pro Xaa
20 25 30

Pro Arg Ser Ile Arg Cys Leu Arg Ala Pro Ala Pro Cys Ser Xaa Pro
35 40 45

Ser Ala Ser Ser Ala Ala Val Gln Val Ala Cys Cys Cys Ser Leu Ala
50 55 60

Cys Cys Gly Pro Ser Arg Pro Ala Ser Gln Gly His Leu Arg Trp Asp
65 70 75 80

Pro Tyr His Leu Ser Arg Asp Leu Tyr Tyr Leu Thr Val Glu Ser Ser
85 90 95

Glu Lys Glu Ser Cys Arg Thr Pro Lys Val Val Asp Ile Pro Thr Tyr
100 105 110

Glu Glu Ala Val Ser Phe Pro Val Ala Glu Gly Pro Pro Thr Pro Pro
115 120 125

Ala Tyr Pro Thr Glu Glu Ala Leu Glu Pro Ser Gly Ser Arg Asp Ala
130 135 140

Leu Leu Ser Thr Gln Pro Ala Trp Pro Pro Pro Ser Tyr Glu Ser Ile
145 150 155 160

Ser Leu Ala Leu Asp Ala Val Ser Ala Glu Thr Thr Pro Ser Ala Thr
165 170 175

Arg Ser Cys Ser Gly Leu Val Gln Thr Ala Arg Gly Gly Ser
180 185 190

<210> 374

<211> 93

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 374

Gly Ser Thr Gly Leu Trp Arg Gly Asp Arg Gly Pro Ile Glu Gly Gly
1 5 10 15

Pro Gly Met Leu Ala Leu Thr Asp His Ser Arg Pro Met Ser Ser Ser
20 25 30

Arg Pro Pro Ala Pro Gln Gln Thr Lys Leu Thr Asp Leu Ser Arg Gly
35 40 45

Leu Gly Pro Ser Gly Thr Gly Tyr Ser Val Xaa Gly Ala Ser Trp Pro
50 55 60

Gly Trp Ala Val Ala Ser Pro Ser Leu His Gln Ala Lys Gln Ser Val
65 70 75 80

Pro Ala Thr Arg Thr Thr Val Pro Leu Thr Val Met Gln
85 90

<210> 375

<211> 27

<212> PRT

<213> Homo sapiens

<400> 375

Gln Trp Phe Trp Trp Pro Gly Arg Ser Ala Ser Leu Gly Gly Ala Lys
1 5 10 15

Gly Met Gln Pro Pro Ser Leu Ala Ser Trp Pro
20 25

<210> 376

<211> 29

<212> PRT

<213> Homo sapiens

<400> 376

Ser	Ser	Ala	Ala	Val	Gln	Val	Ala	Cys	Cys	Cys	Ser	Leu	Ala	Cys	Cys
1				5				10						15	

Gly	Pro	Ser	Arg	Pro	Ala	Ser	Gln	Gly	His	Leu	Arg	Trp			
				20				25							

<210> 377

<211> 32

<212> PRT

<213> Homo sapiens

<400> 377

Val	Ser	Phe	Pro	Val	Ala	Glu	Gly	Pro	Pro	Thr	Pro	Pro	Ala	Tyr	Pro
1				5				10					15		

Thr	Glu	Glu	Ala	Leu	Glu	Pro	Ser	Gly	Ser	Arg	Asp	Ala	Leu	Leu	Ser
				20				25					30		

<210> 378

<211> 26

<212> PRT

<213> Homo sapiens

<400> 378

Arg	Val	Ser	Phe	Pro	Val	Ala	Glu	Gly	Pro	Pro	Thr	Pro	Pro	Ala	Tyr
1					5				10					15	

Pro	Thr	Glu	Glu	Ala	Leu	Glu	Pro	Ser	Gly						
				20				25							

<210> 379

<211> 95

<212> PRT

<213> Homo sapiens

<400> 379

Ser	Asn	Glu	Ile	Leu	Leu	Ser	Phe	Pro	Gln	Asn	Tyr	Tyr	Ile	Gln	Trp
1					5				10				15		

Leu	Asn	Gly	Ser	Leu	Ile	His	Gly	Leu	Trp	Asn	Leu	Ala	Ser	Leu	Phe
					20			25					30		

Ser	Asn	Leu	Cys	Leu	Phe	Val	Leu	Met	Pro	Phe	Ala	Phe	Phe	Leu	
					35			40				45			

Glu	Ser	Glu	Gly	Phe	Ala	Gly	Leu	Lys	Lys	Gly	Ile	Arg	Ala	Arg	Ile
					50			55			60				

Leu	Glu	Thr	Leu	Val	Met	Leu	Leu	Leu	Ala	Leu	Ile	Leu	Gly		
					65			70			75		80		

Ile Val Trp Val Ala Ser Ala Leu Ile Asp Asn Asp Ala Ala Ser
85 90 95

<210> 380

<211> 33

<212> PRT

<213> Homo sapiens

<400> 380

Pro Thr Arg Pro Val Leu Leu Leu Ala Ile Asn Gly Val Thr Glu Cys
1 5 10 15

Phe Thr Phe Ala Ala Met Ser Lys Glu Glu Val Asp Arg Tyr Asn Phe
20 25 30

Val

<210> 381

<211> 93

<212> PRT

<213> Homo sapiens

<400> 381

Asn Asp Lys Lys Leu Leu Phe Leu Lys Gly Phe Trp Ser Ser Leu Lys
1 5 10 15

Asn Glu Thr Pro Pro Pro His Phe Arg Leu Arg Met Val Thr Gly Val
20 25 30

Ser Cys Ser Gly Thr Leu Trp Cys Leu Ile Ser Gly Val Ala Val Thr
35 40 45

Pro Leu Gln Ser Pro Gln Trp Gly Ser Tyr Thr Glu Cys Val Pro Pro
50 55 60

Thr Glu Leu Pro Ile Ala Gly Pro Gly Ala Ser Gly Val Gln Ala Ser
65 70 75 80

Leu Lys Ser Arg His Phe Val Ser Ala Ser Gly His Thr
85 90

<210> 382

<211> 65

<212> PRT

<213> Homo sapiens

<400> 382

Ser Glu Asn Arg Ile Tyr Arg Asn Gly Leu Glu Lys Met Arg Arg Glu
1 5 10 15

Val Thr Ile Gly Arg Ser Ser Ser Ile Cys Leu Asp Gln Gln Val Lys
20 25 30

Ala Gly Asn Ala Val His His Gln Trp Leu Lys Tyr Val Cys Trp Met
35 40 45

Val Val Val Val Gly Gly Ser Gly Val Gly Asp Gly Gly Asn Leu Gly
50 55 60

Met
65

<210> 383
<211> 129
<212> PRT
<213> Homo sapiens

<400> 383
Asn Trp Ser Gly Arg Arg Leu Arg Met Trp Pro Ser Ala Ala Leu Ser
1 5 10 15

Pro Ala Val Ser Ser Pro Ala Leu Ala Leu Thr Ser Pro Pro Lys Pro
20 25 30

Leu Lys Gly Glu Val Trp Leu Arg Trp Lys Leu Leu Gly Ser Arg Ala
35 40 45

Val Gly Leu Phe Ala Phe Ile Ala Leu Gly Thr Gln Ser Pro Leu Leu
50 55 60

His Arg Ala Cys Leu Pro Val Arg Gln Ser Trp Gly Cys Ser Glu His
65 70 75 80

Lys Ala Tyr Pro Ile Leu Arg Leu Gln Pro Asp Leu Glu Thr Gln Val
85 90 95

Gly Pro Gly His Gly Val Asn Trp Asp Leu Arg Thr Gln Ile Arg Thr
100 105 110

Ile Gly Glu Leu Gly Gly Asp Gly Gly Cys Ser Glu Met Arg Pro Leu
115 120 125

Phe

<210> 384
<211> 123
<212> PRT
<213> Homo sapiens

<400> 384
Asn Leu Phe Ser Thr Pro Cys Lys Arg Gln Lys Leu Ile Lys Leu Glu
1 5 10 15

Trp Thr Glu Ala Pro Asn Val Ala Leu Arg Cys Ser Leu Ser Cys Ser
20 25 30

Leu Ile Pro Gly Leu Ser Pro Asp Leu Ser Ser Glu Ala Pro Glu Gly
35 40 45